

ELECKART - MARKETING MIX MODEL

UPGRAD IIIT B Data Science Ecommerce Capstone

Yatin Kode
Sneha Bhatt
Shambhavi Verma
Rohit Behera





Objective:

Develop a market mix model to observe the actual impact of different marketing variables and recommend the optimal budget allocation for the next year.



What is MMM?

- Technique which helps in quantifying the impact of several marketing inputs on sales or Market Share.
- MMM is used to understand how much each marketing input contributes to sales, and how much to spend on each marketing input.

How it works?

MMM uses the Regression technique and the analysis performed through Regression for extracting key information/insights.

OUR BIG IDEA(Approach)



1. Five different ways of predictive modeling is used to determine the most important marketing levers.

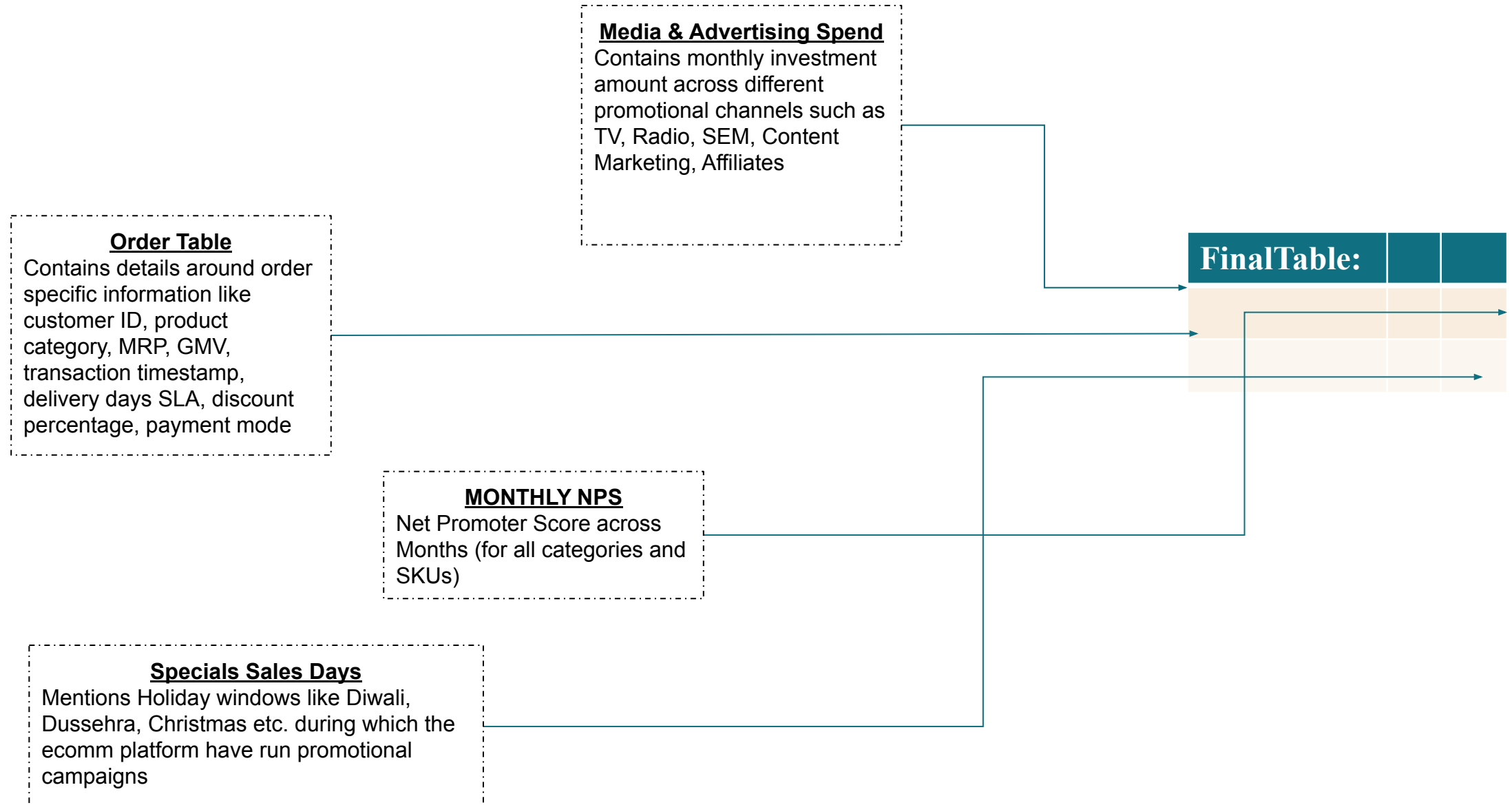


2. The collective data caters to
 - Order Level Data
 - Monthly spends on advertising channels
 - Special sale Days
 - Monthly NPS score

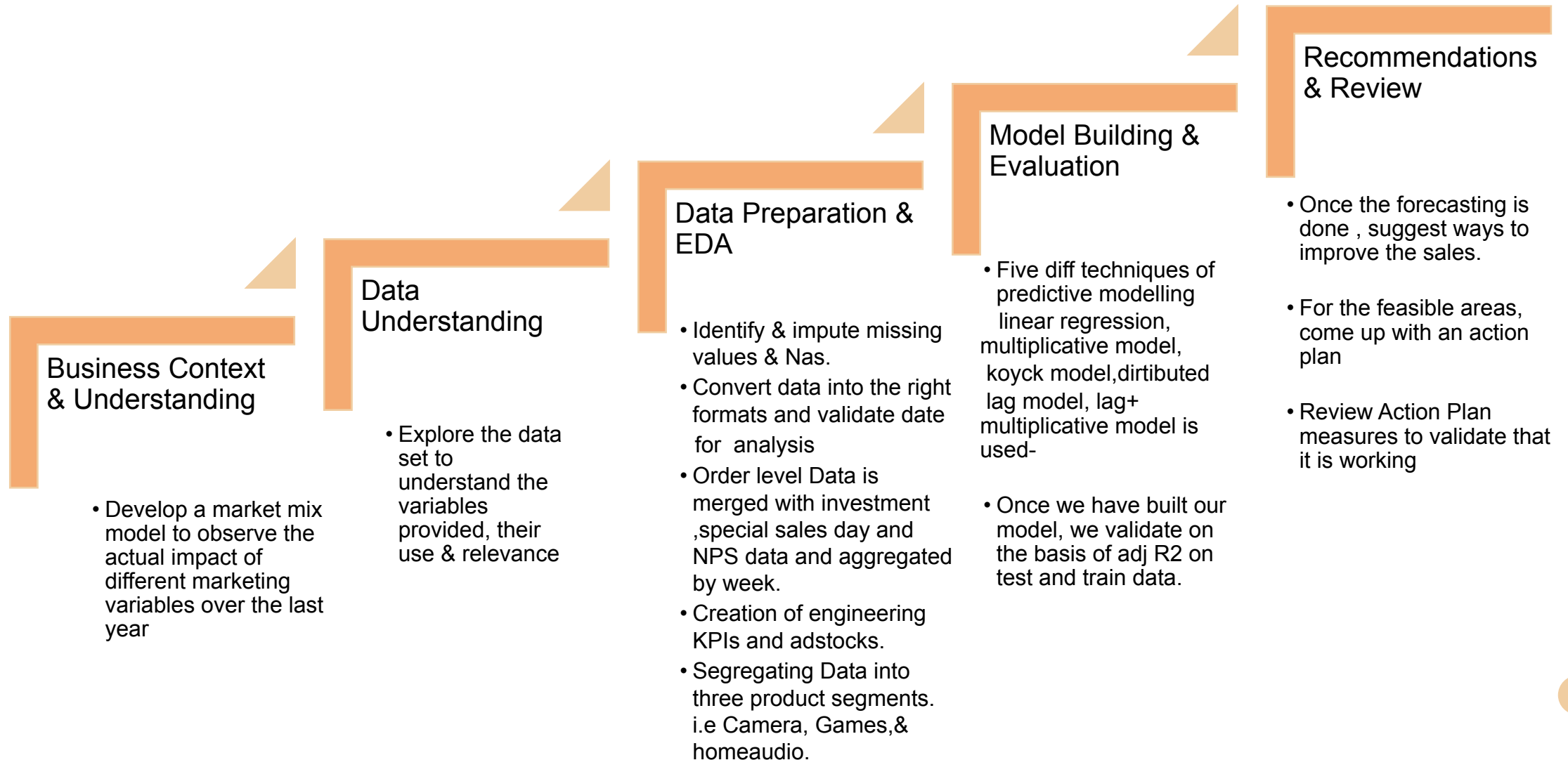


3. Adjusted R square and elasticity coefficients are used to predict most important variables.

The Data Summary



The Data Journey



Data Preparation

1. The data currently has the transaction level data, where each row represents a particular order made from July 2015 to June 2016
2. There are 12 attributes related to order level data.
3. Scope of analysis is limited to three product categories - Home Audio, Gaming Accessories and Camera Accessories
4. The data is aggregated **on weekly basis** and merged with NPS, Media Investments and Special Sales Data.
5. Since the independent attributes had a widely varying range of value, they were scaled prior to modeling step.



Feature Engineering

KPI	Formula	Rationale
Discount over MRP	MPP - gmv	Discount rate is one of the important factor effecting sales
Holiday week	Flag indicating whether there were holidays in the week or it was a regular week. (Based on Sales Calendar Data)	Segregating special sale days to understand effect sales on Normal Days vs Holidays
Payment Mode	Prepaid or COD	Customer payment preference can be an important revenue driver. Ecommerce platform offering the maximum number of payment options especially COD for high value items may attract more traffic.
AdStock for each of the promotional channels	Adstock describe the prolonged or lagged effect of advertising on consumer purchase behavior. Adstock variables are created for TV, Radio, Content marketing, Digital, Sponsorship, SEM and Others	We need to consider the lagged effect of all marketing channels, to capture the after effect of the investment done
Lag variables for price based variables (GMV and list price)	Weekly lag in price is captured in lag variables, basically needed for lag based predictive models	Needed for Lag based predictive modelling

- Now since the target variable is decided , three separate data frames are created for Home, Audio and Gaming accessories
- Important engineering Kpis – List Price, Promotional Offer , Discount over MRP, Holiday week, No promotion week are derived
- Discount range, Promotion type, Payment type, Delivery status are important derived matrices for analysis.



Some important terms:

Models used:

1. Multi-Linear Regression:

where $x_1, x_2, \dots, X(n)$ are Market Predictors..

$$\text{Sales} = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2$$

Sales Equation

2. Multiplicative Model:

$$S_t = \exp(\alpha_0) \cdot (P_t^{\alpha_1}) \cdot (A_t^{\alpha_2}) \cdot \exp(\epsilon_t) \quad (5)$$

Where S_t =sale, a_1, a_2, \dots, a_n are Market predictors...

3. Koyck Model:

$$S_t = \beta_0 + \sum_{k=1}^K \beta_k X_{kt} + \lambda S_{t-1} + v_t$$

Where S_t =Sale and $\lambda S_{t-1} + v_t$ is the carry over effect on sales due to promotion.

Let's talk some more Technical:

Models used:

4. Distributed Lag Model:

$$y_t = a + w_0x_t + w_1x_{t-1} + w_2x_{t-2} + \dots + \text{error term}$$

where current values of a dependent variable = current values of an market predictors and the lagged (past period) values of these variable.

5. Distributed Multiplicative Lag Model

$$St = \exp(\alpha_0) \cdot (Pt^{\alpha_1}) \cdot (At^{\alpha_2}) \cdot \exp(\epsilon t) + y_t = a + w_0x_t + w_1x_{t-1} + w_2x_{t-2} + \dots + \text{error term}$$

Where The multiplicative effect of predictors and their lag values are also considered while predicting sales.



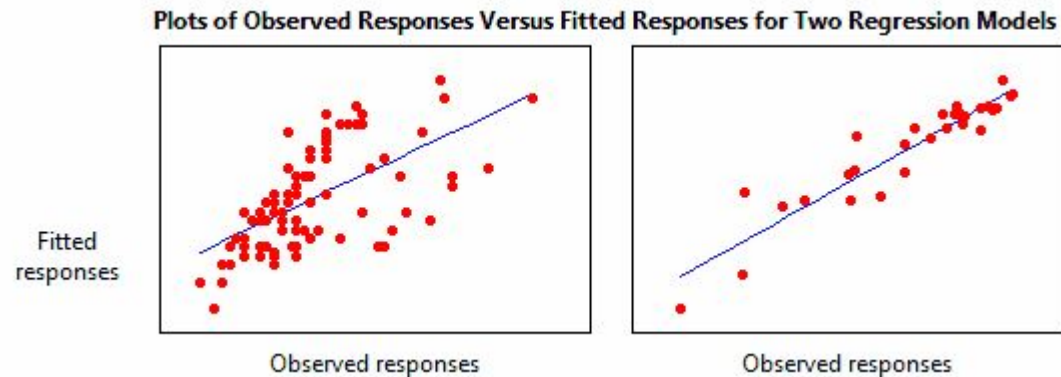
Let's talk some more Technical:

Evaluation Matrices used:

1. Adjusted R square:

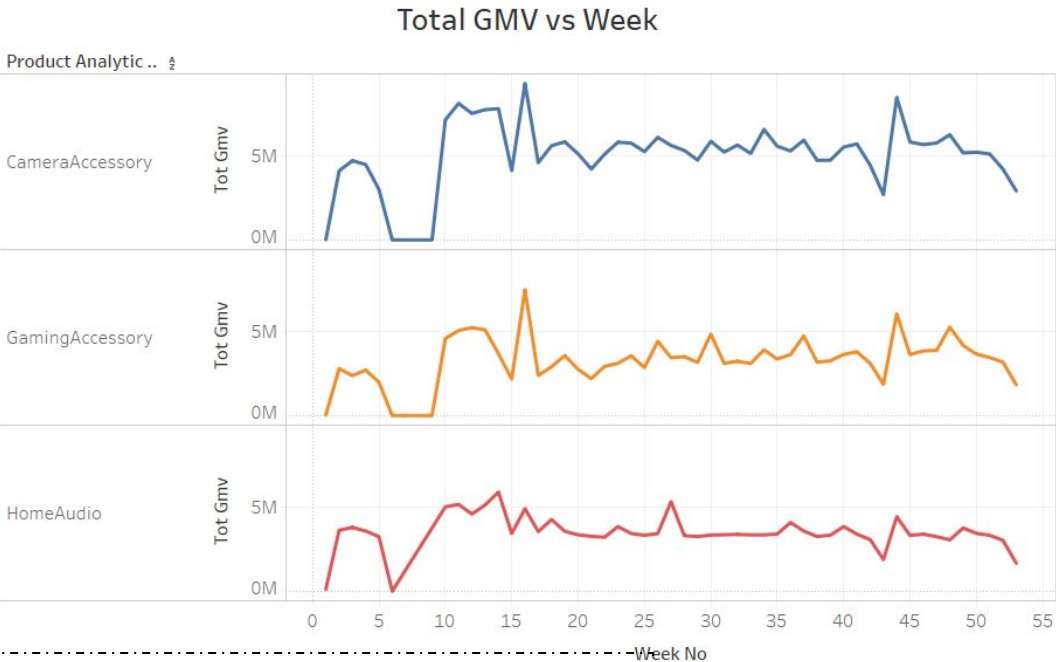
Tells how close is the data to the fitted regression line. Higher the value, Accurate is the model.

R^2 is good for second image where the data points are closely aligned.

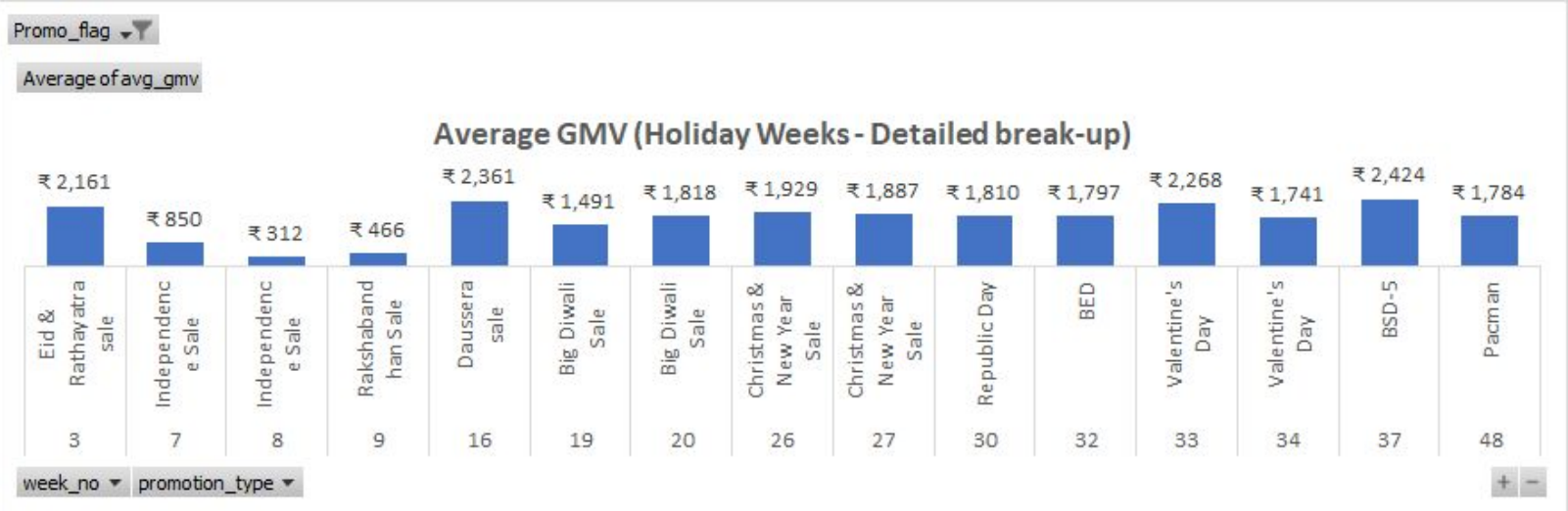
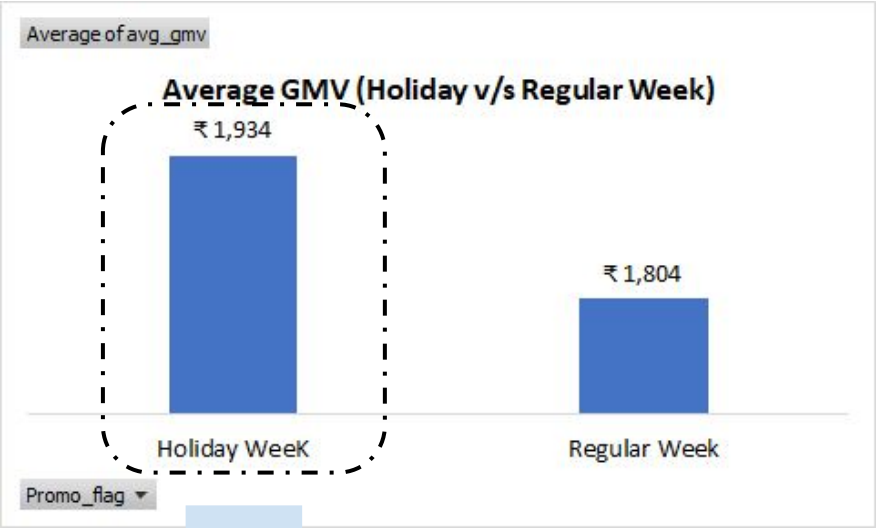


Product Wise Analysis and Model Interpretation and Evaluation:

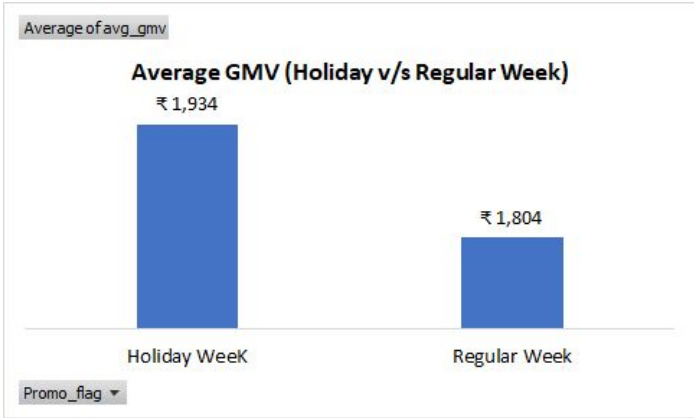
Average GMV on a regular week v/s Special Holiday week (All Product Categories)



- Average GMV during Holiday Week is 7% more than in the Regular Week. This suggests holiday weeks are more profitable
- Among holiday week, Christmas, New Year, Pacman and BDS-5 shows more GMV
- There is a sharp spike during the weeks between 10 to 15, in the event of Dussehra and Diwali.
- Subtle spikes during the promotional periods across the year.



Average GMV on a regular week v/s Special Holiday week (All Product Categories)



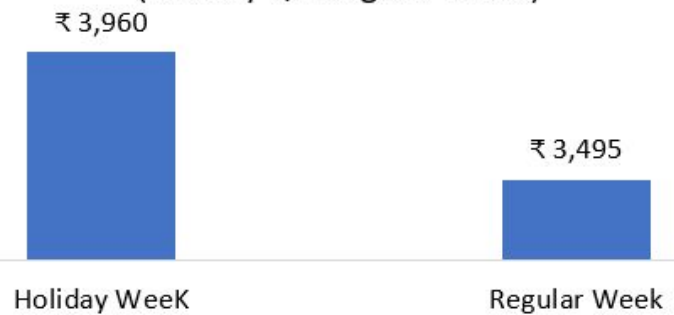
Camera Accessories - Avg GMV
(Holiday v/s Regular Week)



Gaming Accessories - Avg GMV
(Holiday v/s Regular Week)



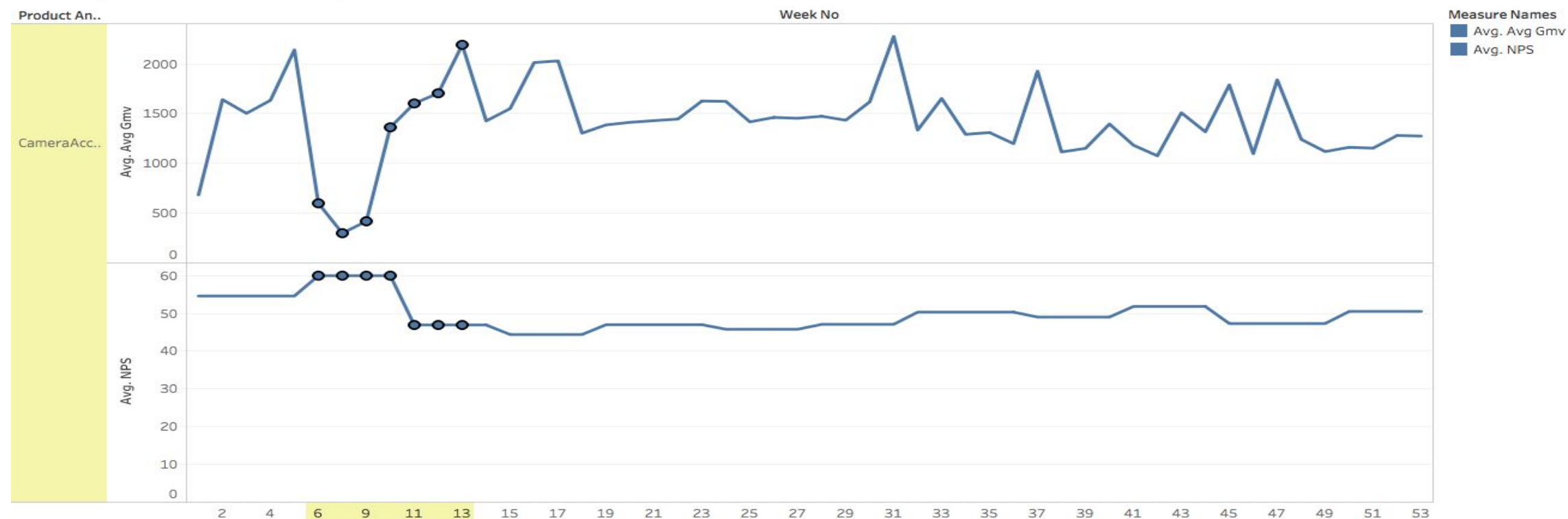
Home Accessories - Avg GMV
(Holiday v/s Regular Week)



Home Accessories category saw maximum decline in average GMV between Regular and Holiday weeks. (~15%) when compared with Gaming Accessories (<1%) and Camera Accessories (~5%)

Lagged Effect of NPS on GMV (Category : Camera Accessories)

<Gmv v/s Disocunt over MRP/Value per customer>



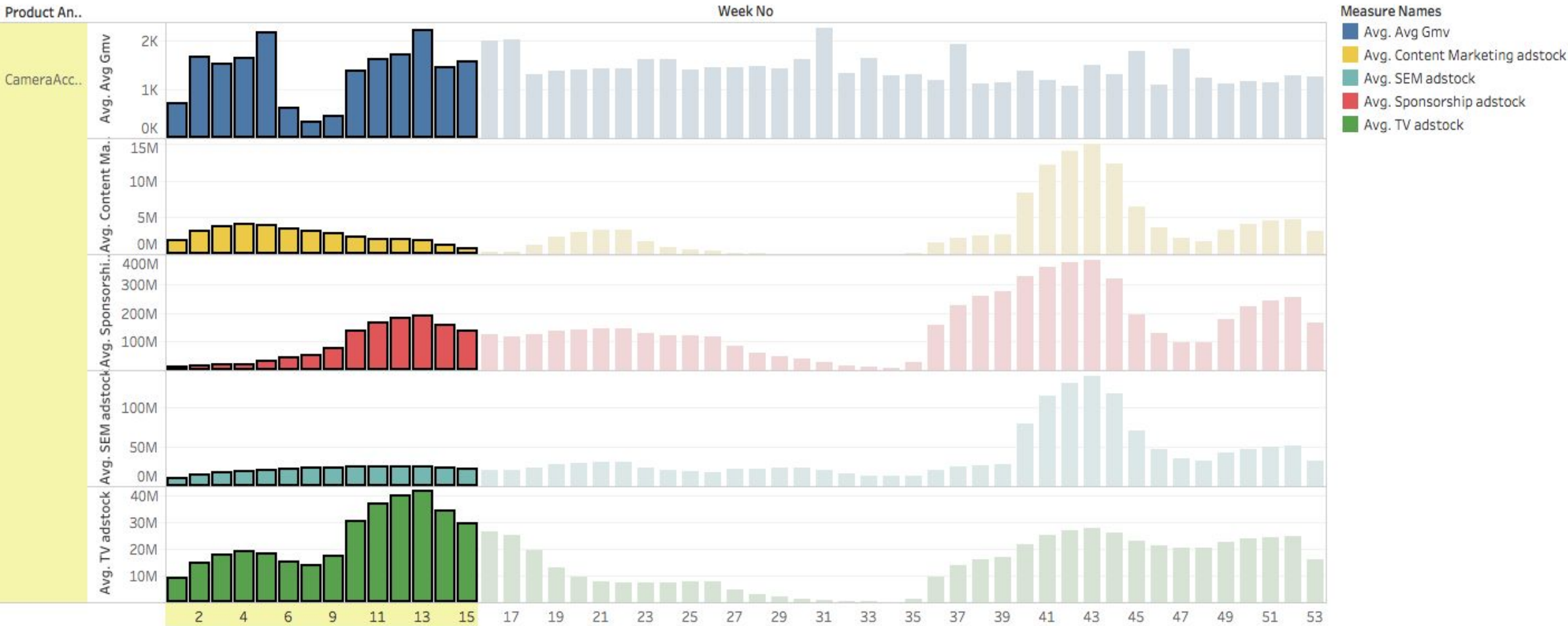
The trends of Avg. Avg Gmv and Avg. NPS for Week No broken down by Product Analytic Sub Category. Color shows details about Avg. Avg Gmv and Avg. NPS. The view is filtered on Product Analytic Sub Category, which keeps CameraAccessory.

The impact of of gmv on NPS score. Higher the gmv(week no.4 to 6),NPS score(6-10) improves successively while as gmv reduces(week no. 6-9) , the NPS score went down(week no. 11-13)



Effect of Different Investments on GMV of Camera Accessories (First few weeks)

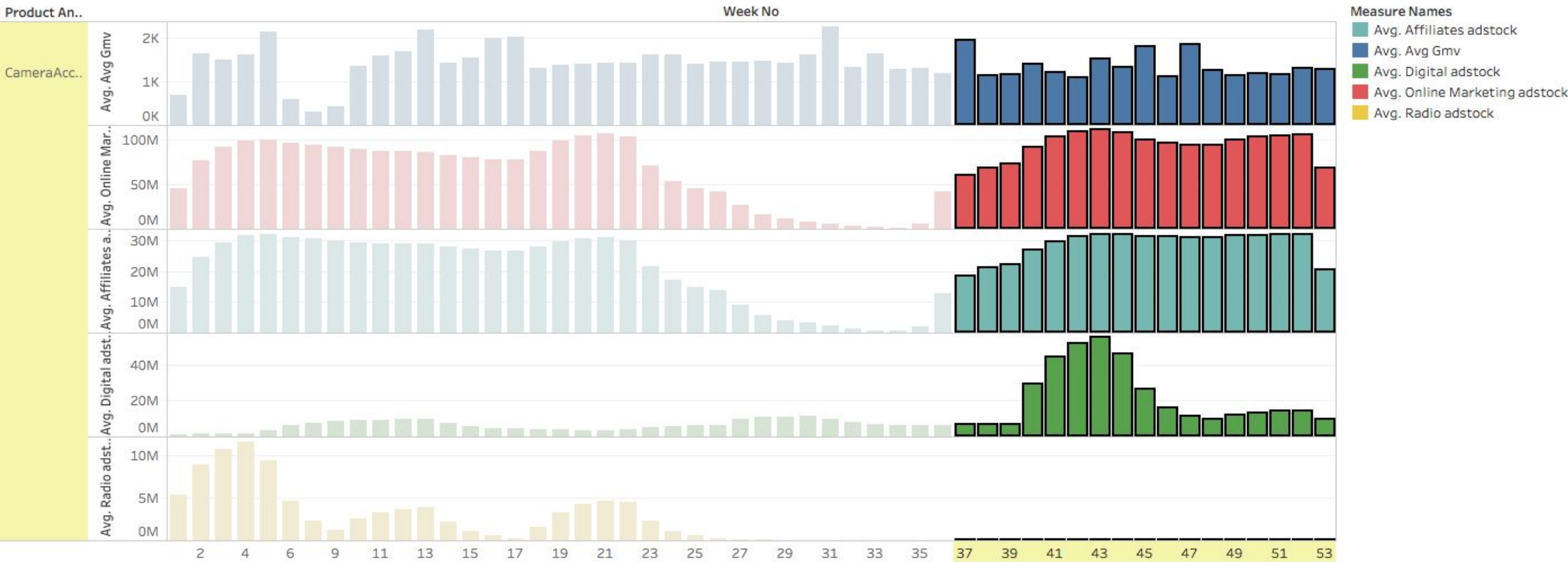
<Gmv v/s Investmnet across adstocks>



More investment in TV/Content Marketing(week no.2 to 6) shows higher GMV (week no.9 to 15)

Effect of Different Investments on GMV of Camera Accessories

<Gmv v/s Promotion type-sale holidays>



No Trend in GMV was observed in the weeks 37 to 53.

Effect of Delivery SLA on GMV of Camera Accessories (Last few weeks)

<Gmv v/s Early/Ontime/LateDelivery>

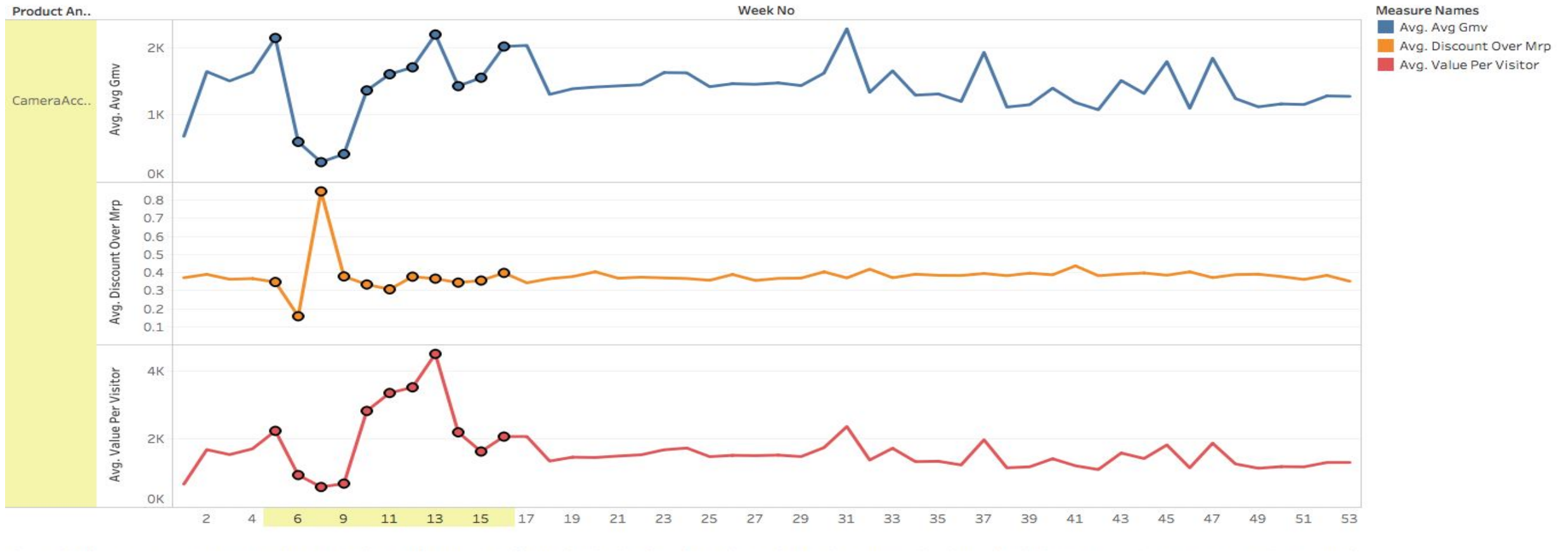


Delivery delays affect GMV for Camera Accessories. With increased delays, GMV goes down.



Effect of Discount on Average GMV per customer

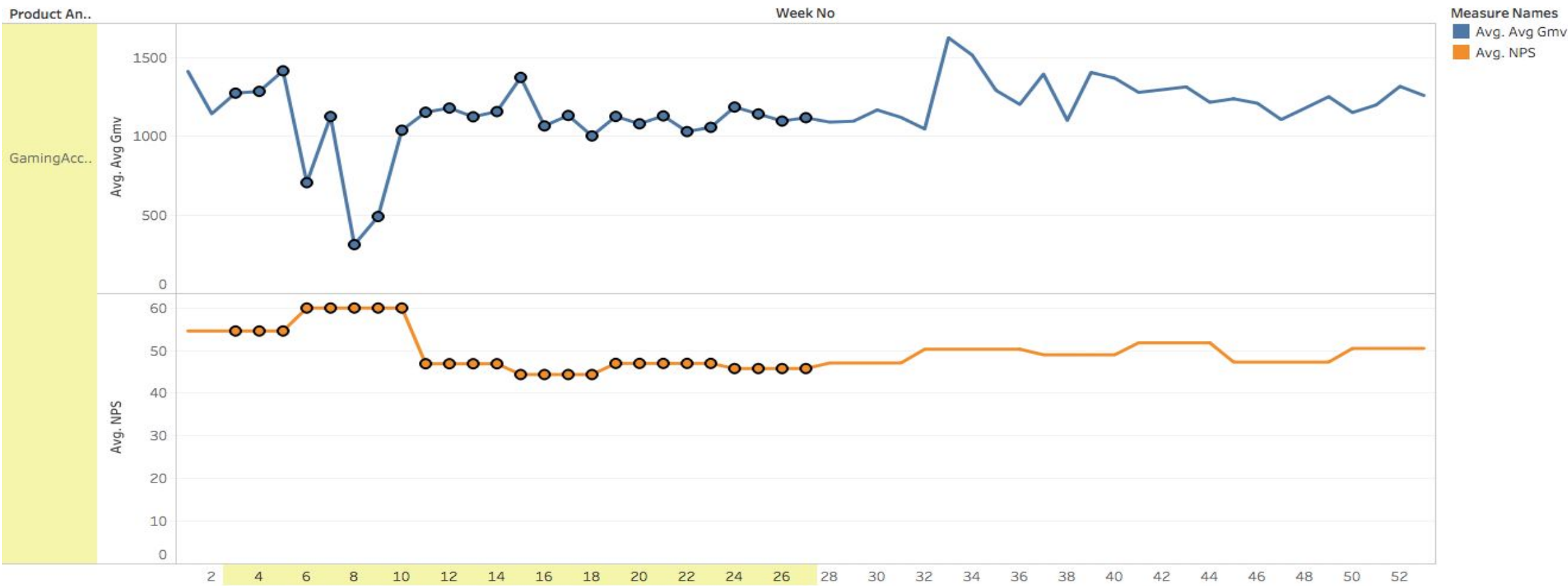
<Gmv v/s Disocunt over MRP/Value per customer>



- Higher Value Per customer aggregates to Higher GMV
- More Discount over MRP, creates more revenue.

Lagged Effect of NPS on GMV (Category : Gaming Accessories)

<Gmv v/s Promotion type-sale holidays>



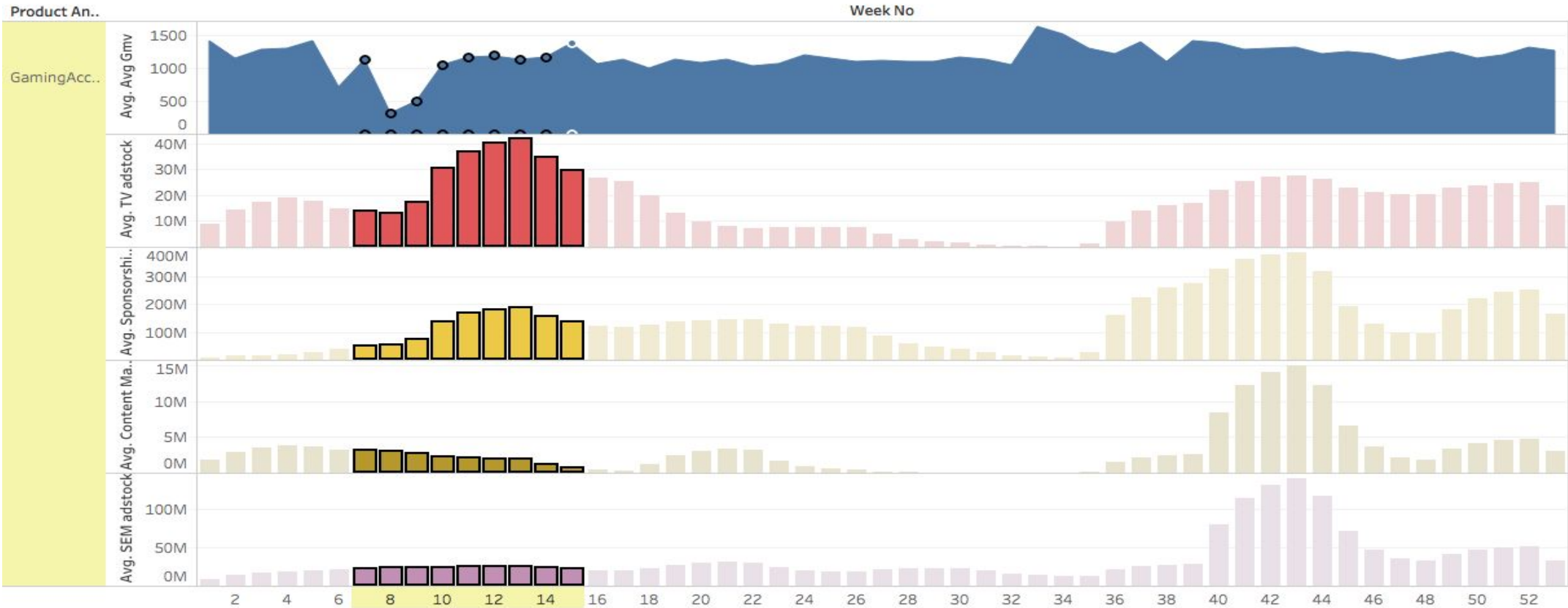
The trends of Avg. Avg Gmv and Avg. NPS for Week No broken down by Product Analytic Sub Category. Color shows details about Avg. Avg Gmv and Avg. NPS. The view is filtered on Product Analytic Sub Category, which keeps GamingAccessory.

Higher the gmv(week no. 3 to 6),NPS score(week no.6 to 8) improves successively while as gmv reduces(week no.6 to 9) , the NPS score went down(week no. 11 to 13)



Effect of Different Investments on GMV of Gaming Accessories (Week# 6 to 15)

<Gmv v/s Investment on diff media investment>



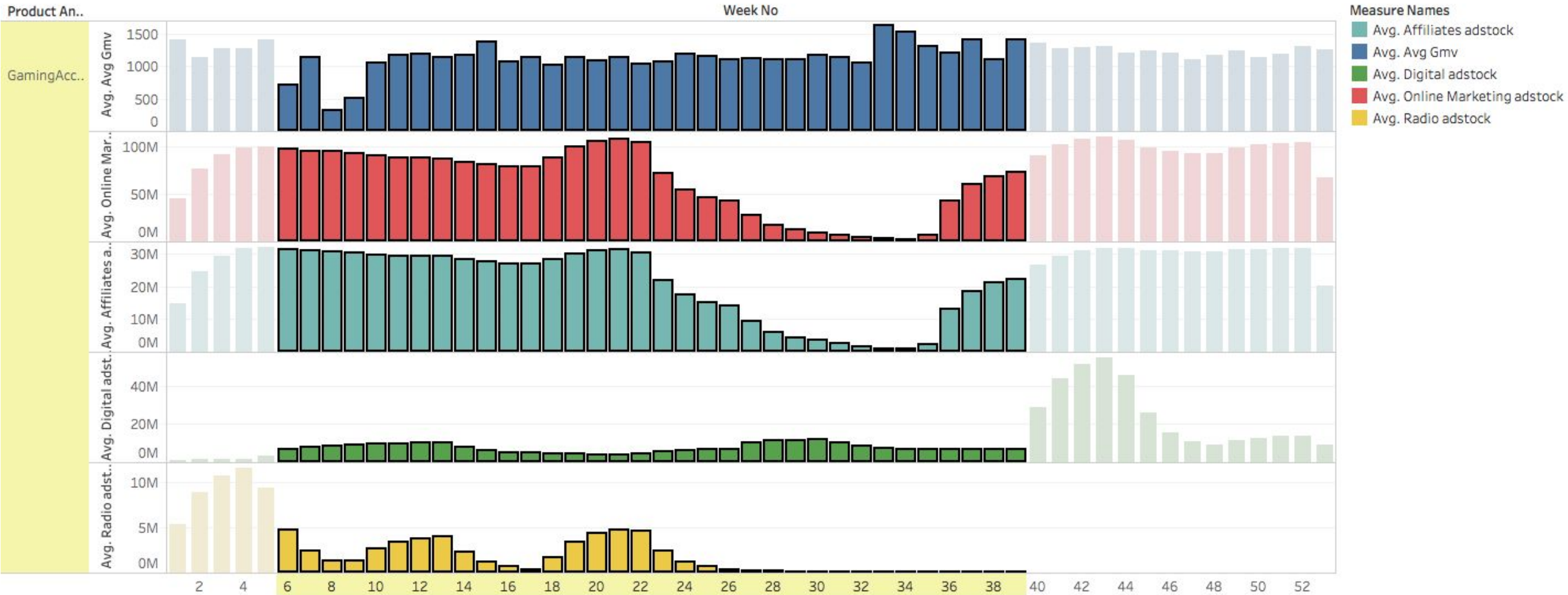
Average of Avg Gmv, average of TV adstock, average of Sponsorship adstock, average of Content Marketing adstock and average of SEM adstock for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps GamingAccessory.

More investment in TV/Content Marketing(week# 6 to 8) shows higher GMV (week# 10-14)



Effect of Different Investments on GMV of Gaming Accessories (Week# 33 to 38)

<Gmv v/s Promotion type-sale holidays>

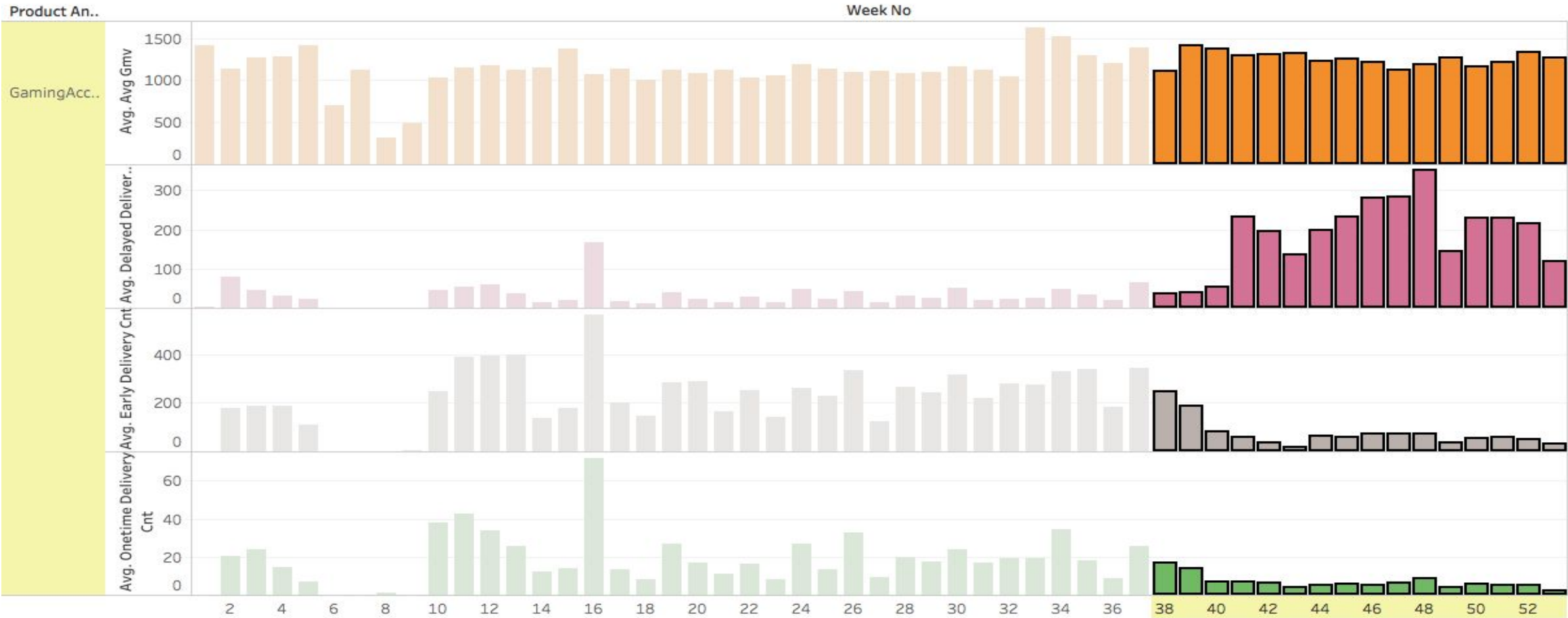


Avg. Avg Gmv, Avg. Online Marketing adstock, Avg. Affiliates adstock, Avg. Digital adstock and Avg. Radio adstock for each Week No broken down by Product Analytic Sub Category. Color shows details about Avg. Avg Gmv, Avg. Online Marketing adstock, Avg. Affiliates adstock, Avg. Digital adstock and Avg. Radio adstock. The view is filtered on Product Analytic Sub Category, which keeps GamingAccessory.

Investments in weeks 18 to 24 leads to slight upsurge in GMV in a lagged manner.

Effect of Delivery SLA on GMV of Gaming Accessories (Last few weeks)

<Gmv v/s Early and Delayed delivery week wise>

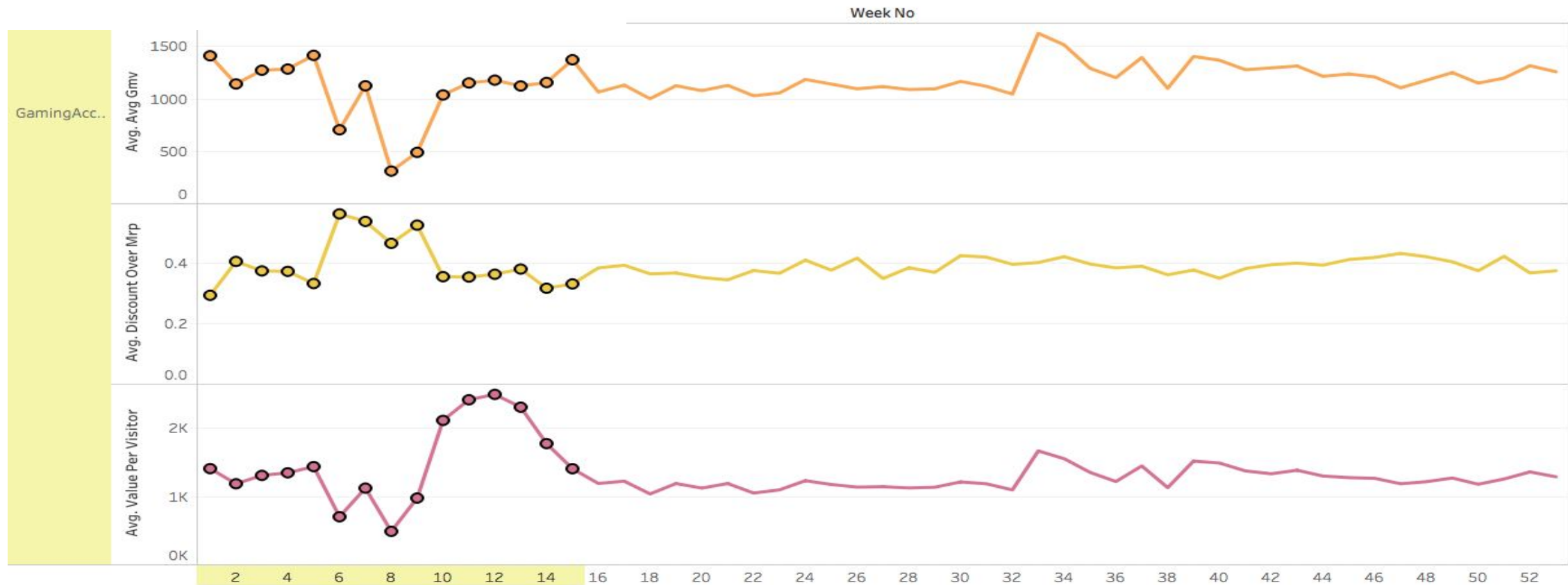


Average of Avg Gmv, average of Delayed Delivery Cnt, average of Early Delivery Cnt and average of Onetime Delivery Cnt for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps GamingAccessory.

- Delivery SLA affect sales.
- Avg sales is going down as delivery is delayed.



Gmv v/s Discount over MRP and Value Per visitor (Gaming Accessories)

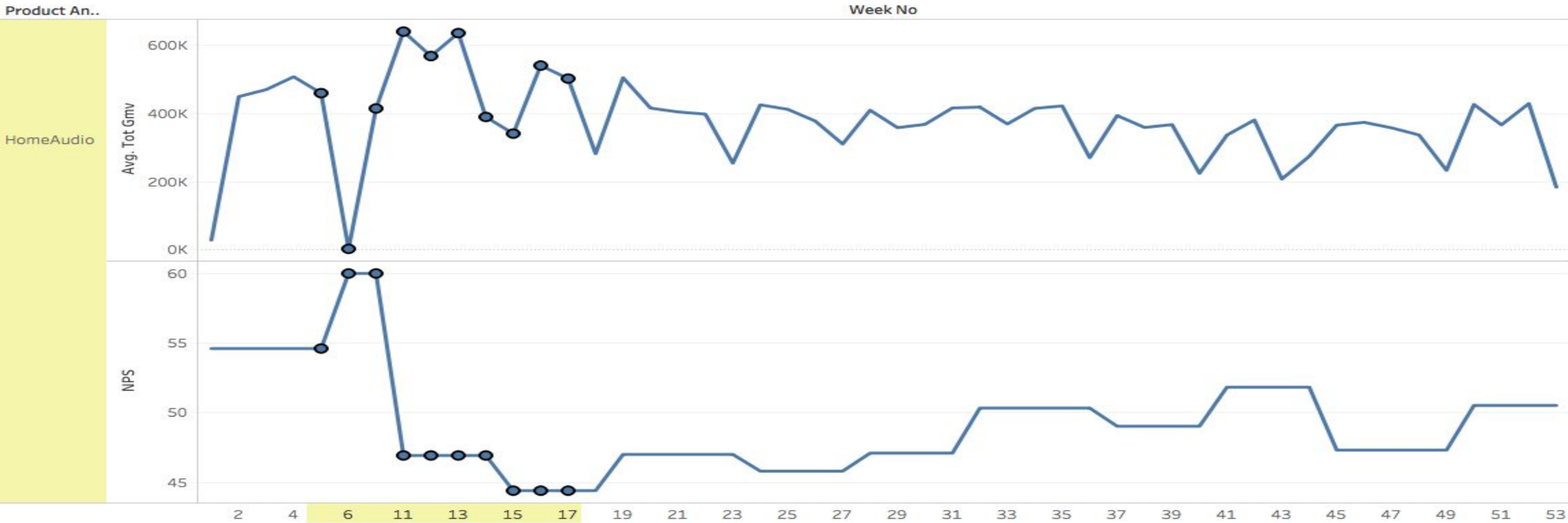


The trends of average of Avg Gmv, average of Discount Over Mrp and average of Value Per Visitor for Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps GamingAccessory.

- Higher Value Per customer aggregates to Higher GMV!
- More Discount over MRP, creates more revenue.

Lagged Effect of NPS on GMV (Category : Home Audio)

<Gmv v/s NPS score>



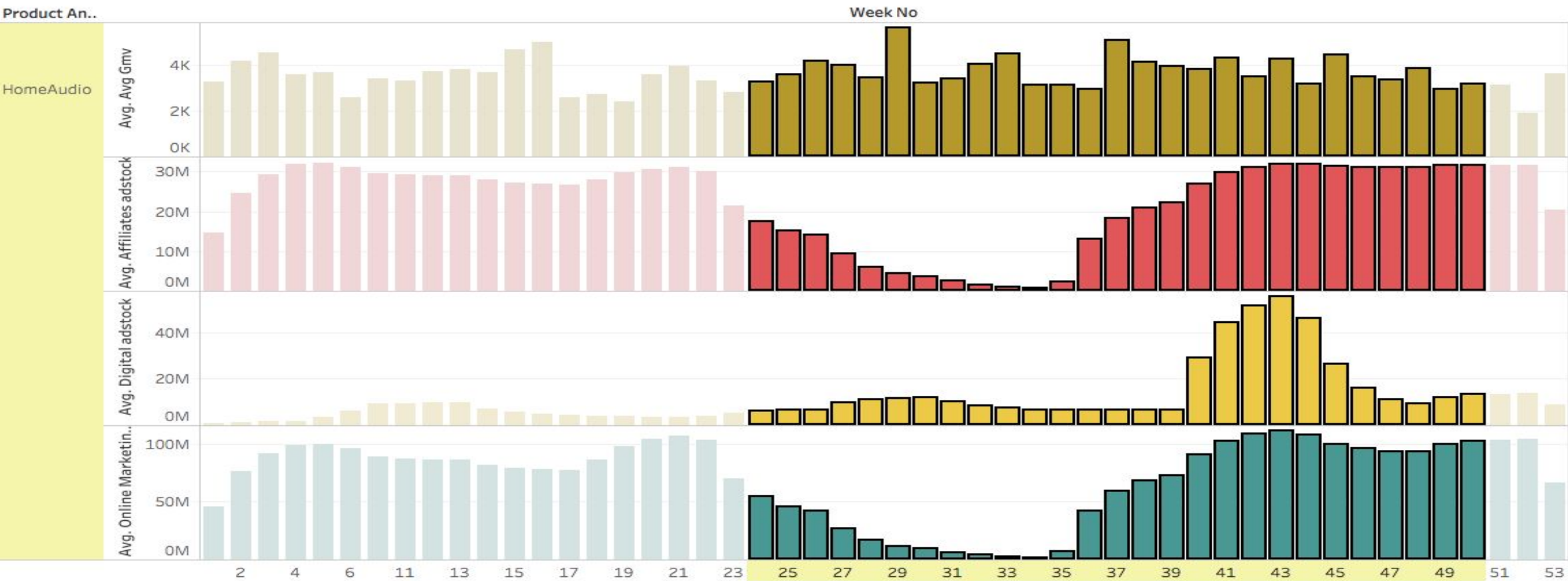
The trends of average of Tot Gmv and NPS for Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps HomeAudio.

The impact of of gmv on NPS score. Lower the gmv(5-7),NPS score(9-11) reduces successively while as gmv is quiet constant(15-...) , the NPS score went constant(20-...)



Lagged Effect of NPS on GMV (Category : Home Audio)

<Gmv v/s Media investment on online channel>



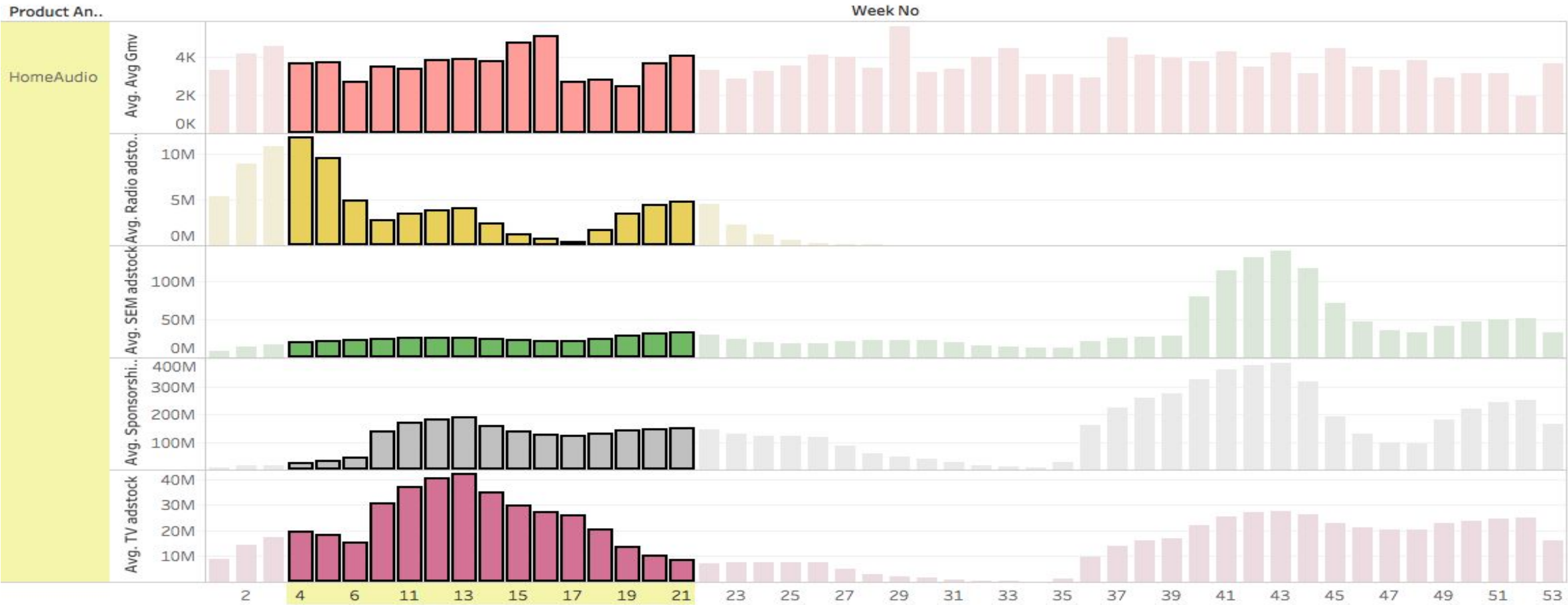
Average of Avg Gmv, average of Affiliates adstock, average of Digital adstock and average of Online Marketing adstock for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps HomeAudio.

Low investment in Affiliates /Online Marketing/ SEM/TV/Sponsorship resulting into lower gmv successively. Less investment in Affiliates/Online Marketing(week no.25 to 33) shows higher gmv (week 36 and beyond..)



Lagged Effect of NPS on GMV (Category : Home Audio)

<Gmv v/s Media investment on online channel>



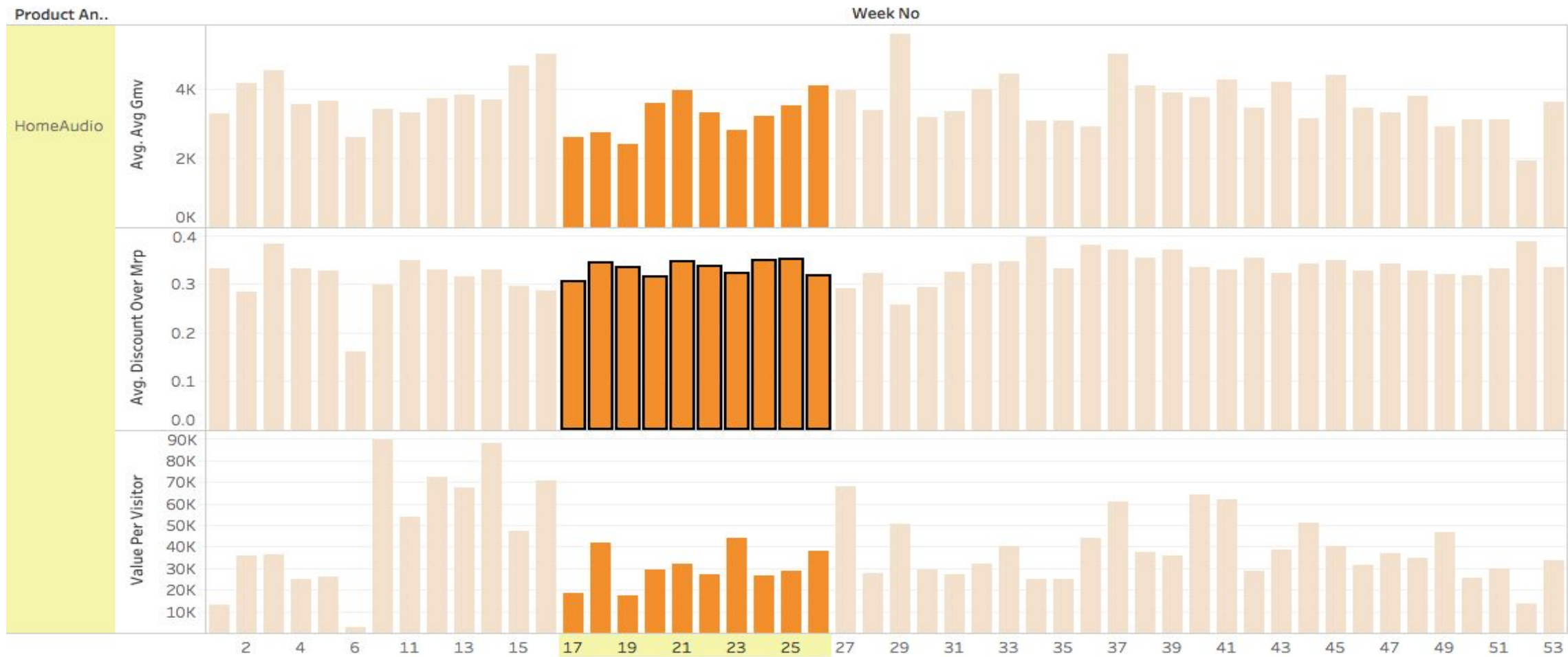
Average of Avg Gmv, average of Radio adstock, average of SEM adstock, average of Sponsorship adstock and average of TV adstock for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps HomeAudio.

Low investment in Affiliates /Online Marketing/ SEM/TV/Sponsorship resulting into lower gmv successively. Less investment in Affiliates/Online Marketing(week no.13 to 21) shows higher gmv (week 23 and beyond..)

Effect of Average Order Value on GMV

Higher Value per customer aggregates to higher GMV. More discounts led to higher revenue, owing significant volume uptick.

<Gmv v/s Early and Delayed delivery week wise>

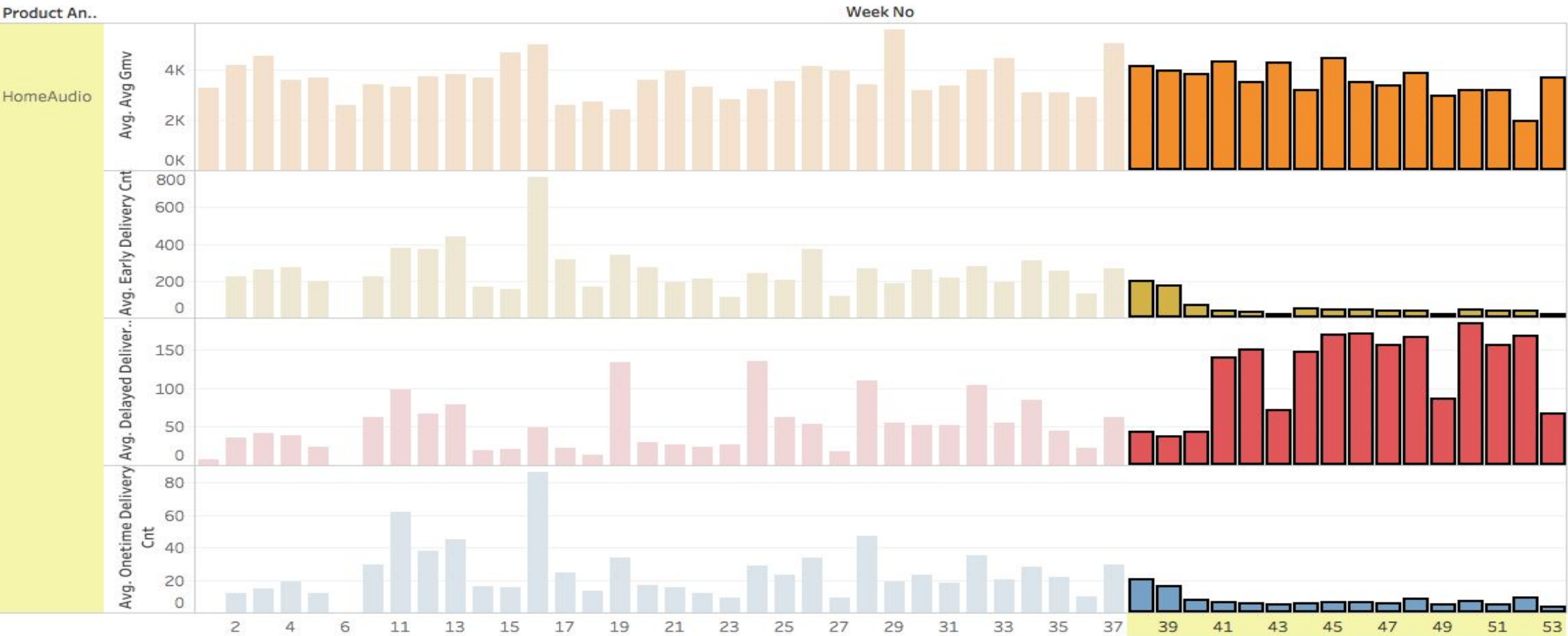


Average of Avg Gmv, average of Discount Over Mrp and sum of Value Per Visitor for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps HomeAudio.

Effect of Delivery Timing on GMV

Delivery delays affect customer experience. Increase in delays for Home Audio category led to decline in weekly GMV.

<Gmv v/s Early/Late delivery>



Average of Avg Gmv, average of Early Delivery Cnt, average of Delayed Delivery Cnt and average of Onetime Delivery Cnt for each Week No broken down by Product Analytic Sub Category. The view is filtered on Product Analytic Sub Category, which keeps HomeAudio.

Data Modeling

Choice of Target Predicted Variable

Target Variable	Rationale
Total GMV (weekly)	This KPI is directly available in the given data set. GMV is unit sale price times number of units. GMV gives a sense of volume and customer traction across categories. Investment across multiple promotional channels affect customer traffic and completion of order placement.
Average Order Size (weekly)	This was a strong contender for Target variable, but eventually dropped off because the variation of order size was minimal across weeks, indicating volumes or GMV are a better fit for target variable
Revenue (weekly)	This was also another strong contender. Revenue is the actual realized cash flow on a transaction between retailer and brand owner. Investment in different promotional channels can lead to volume increase but also opens up the scope of re-negotiation of commission fees. This is commonly observed dynamics in food delivery tech. And there is no reason why E-comm will be any different. Considering revenue opens up additional complexities of widely differing commission percentages across multiple categories and SKUs. To avoid complexities, we abandoned Revenue (weekly) as the target predicted variable. As a simplifying assumption, we have considered the effect of investments across multiple promotional channels drive customer traffic and order volumes.

All subsequent models are built using dependent variable as Total GMV instead of the other two KPIs.

Camera Accessory Model Performance

Model Type	Significant Variables	Training Adjusted R-squared(2)	Testing Adjusted R-squared(2)
Linear Model	<ul style="list-style-type: none"> • Investment Amount (Sponsorship, Content Marketing) • (1)Product Analytic Vertical (Accessory, Battery Charger, Battery Grip, Eye Cup,Film Rolls, Housing,LED Light, Microphone, Mount, Remote Control, Extension Tube, Filter,Lens, Reflector Umbrella,Softbox, Strap, Telescope) • (3)Special Time Windows (Christmas.& New.Year.Sale, Dussehra.sale, No Sale) 	80.9%	69.9%
Multiplicative model	<ul style="list-style-type: none"> • List price, week number, Radio adstock(4), Discount • (1)Product Analytic Vertical (Accessory ,Battery ,Battery Grip ,EyeCup ,Housing ,LEDLight ,Microphone,Mount ,Remote Control ,Extension Tube ,Flash ,Flash Shoe Adapter ,Lens ,Reflector Umbrella ,Softbox ,Strap, Teleconverter ,Telescope) • (3)Special Time Windows (Independence.Sale ,No Sale) 	82.8%	79%
Distributed Lag Model	<ul style="list-style-type: none"> • (1)Product Analytic Vertical (Lens), 1st lag of GMV • (3)Special Time Windows (Dussehra sale, Rakshabandhan Sale) 	69%	61.8%
Koyck Model	<ul style="list-style-type: none"> • week • (1)Product Analytic Vertical (Accessory, Battery Charger, Battery Grip, EyeCup,Film Rolls, Housing,LED Light, Microphone,Mount, Remote Control, Extension Tube, Filter, Lens,Softbox,Strap ,Telescope) • (3)Special Time Windows (Dussehra sale, Rakshabandhan Sale) 	79.4%	73.4%
Lag plus Multiplicative model	<ul style="list-style-type: none"> • Discount • Lag Variables (1st lag of GMV) • (1)Product Analytic Vertical (Lens) • (3)Special Time Windows (Dussehra sale, Rakshabandhan Sale) 	69.7%	53.4%

(1) Variables of Product Analytic Vertical type indicates high GMV opportunity. Hence all marketing/.promotional efforts should be directed towards these SKUs.

(2) R-squared values indicate the percentage of variation explained by the obtained value, i.e Linear model explains 80.9% of the variation in training set and 69.9% of variation in the test set.

(3) Spike in GMV was noticed in Christmas and New Year for Linear Model. Hence promotional budget should be allocated for these special periods for higher revenue realization.

(4) Advertising adstock describes the prolonged or lagged effect of advertising on consumer purchase behavior. ... Adstock is a model of how response to advertising builds and decays in consumer markets.

Game Accessory Model Performance

Model Type	Significant Variables	Training Adjusted R-squared(2)	Testing Adjusted R-squared(2)
Linear Model	<ul style="list-style-type: none"> • Investment Amount (TV, Affiliates, Radio) • (1)Product Analytic Vertical (GamePad, Gaming Headset, Gaming Keyboard, Gaming Mouse) • (3)Special Time Windows (Christmas.& New.Year.Sale, Rakshabandhan Sale, No Sale) 	72.59%	62%
Multiplicative model	<ul style="list-style-type: none"> • List price, week, NPS • Investment Amount (Total Investment, TV, Affiliates, SEM , Other) • (4)Advertising Adstock (Sponsorship adstock, SEM adstock) • (1)Product Analytic Vertical (Cooling Pad , Game Control Mount , GamePad , Gaming Adapter , Gaming Charging Station , Gaming Gun , Gaming Headset , Gaming Mouse , Gaming Speaker , Joystick Gaming Wheel , Motion Controller) • (3)Special Time Windows (Rakshabandhan Sale ,No Sale) 	76.8%	76.6%
Distributed Lag Model	<ul style="list-style-type: none"> • Week, Other Investment,, first weeklag of GMV • (1)Product Analytic Vertical (GamePad, Gaming Headset, Gaming Mouse) • (3)Special Time Windows (Dussehra sale, Independence Sale, Rakshabandhan Sale) 	73%	68.8%
Koyck Model	<ul style="list-style-type: none"> • Investment Amount (TV) • (1)Product Analytic Vertical (GamePad, Gaming Headset, Gaming Keyboard, Gaming Mouse) • (3)Special Time Windows (No Sale, Rakshabandhan Sale) 	89.1%	70.8%
Lag plus Multiplicative model	<ul style="list-style-type: none"> • List price, NPS, TV adstock(4), Radio Adstock(4), 2nd and 3rd lag of GMV, 1st week lag of GMV, 1st lag of List Price • Investment Amount (Total Investment, TV, Content Marketing, Other) • (1)Product Analytic Vertical (Cooling Pad, GamePad, Gaming Accessory Kit, Gaming Charging Station , Gaming Headset, Gaming Keyboard, Gaming Mouse, Gaming Speaker) • (3)Special Time Windows (Independence Sale, Rakshabandhan Sale, No Sale) 	81.3%	77.6%

(1) Variables of Product Analytic Vertical type indicates high GMV opportunity. Hence all marketing/.promotional efforts should be directed towards these SKUs.

(2) R-squared values indicate the percentage of variation explained by the obtained value, i.e Linear model explains 80.9% of the variation in training set and 69.9% of variation in the test set.

(3) Spike in GMV was noticed in Christmas and New Year for Linear Model. Hence promotional budget should be allocated for these special periods for higher revenue realization.

(4) Advertising adstock describes the prolonged or lagged effect of advertising on consumer purchase behavior. ... Adstock is a model of how response to advertising builds and decays in consumer markets.

Home Audio Model Performance

Model Type	Significant Variables	Training Adjusted R-squared(2)	Testing Adjusted R-squared(2)
Linear Model	<ul style="list-style-type: none"> • Investment Amount (Sponsorship, Content Marketing) • (1)Product Analytic Vertical (FM Radio, Home Audio Speaker, Voice Recorder) 	86.6%	76.7%
Multiplicative model	<ul style="list-style-type: none"> • List price, Discount • Investment Amount (Other) • (1)Product Analytic Vertical (DJ Controller, Docking Station, FM Radio, HiFi System, Home Audio Speaker, Karaoke Player, SlingBox, Sound Mixer) 	79.5%	64%
Distributed Lag Model	<ul style="list-style-type: none"> • Lag Variables (3rd lag of GMV, 1st week lag of GMV) • (1)Product Analytic Vertical (Home Audio Speaker) 	78.1%	72.1%
Koyck Model	<ul style="list-style-type: none"> • (1)Product Analytic Vertical (FM Radio, Home Audio Speaker) 	89.1%	70.8%
Lag plus Multiplicative model	<ul style="list-style-type: none"> • List price, Discount, Affiliated Adstock(4) • Investment Amount (Radio) • Lag Variables (2nd and 3rd lag of GMV) • (1)Product Analytic Vertical (DJ Controller, Dock, Docking Station, FM Radio, HiFi System, Home Audio Speaker, Sound Mixer) 	84.1%	84.2%

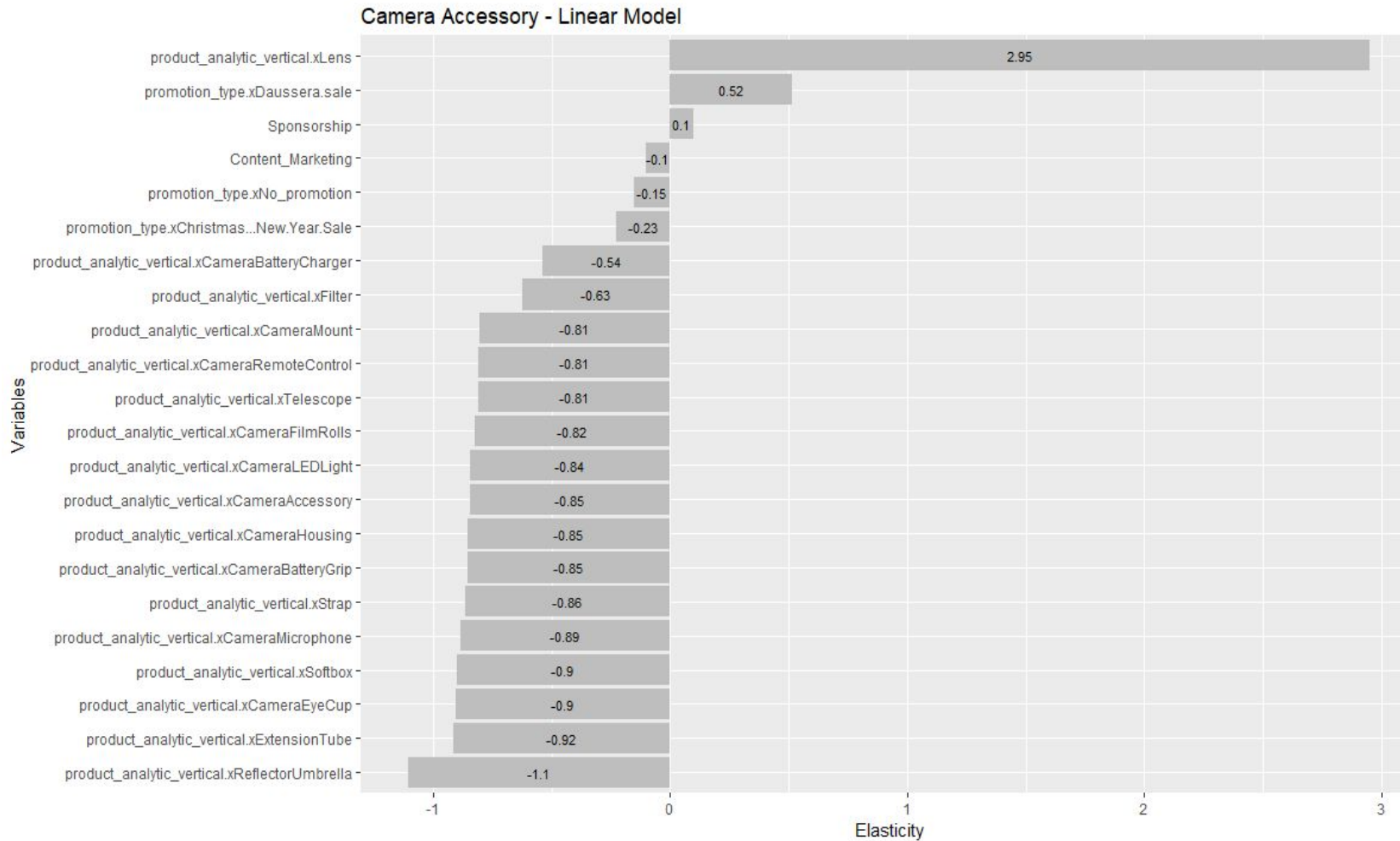
(1) Variables of Product Analytic Vertical type indicates high GMV opportunity. Hence all marketing/.promotional efforts should be directed towards these SKUs.

(2) R-squared values indicate the percentage of variation explained by the obtained value, i.e Linear model explains 80.9% of the variation in training set and 69.9% of variation in the test set.

(3) Spike in GMV was noticed in Christmas and New Year for Linear Model. Hence promotional budget should be allocated for these special periods for higher revenue realization.

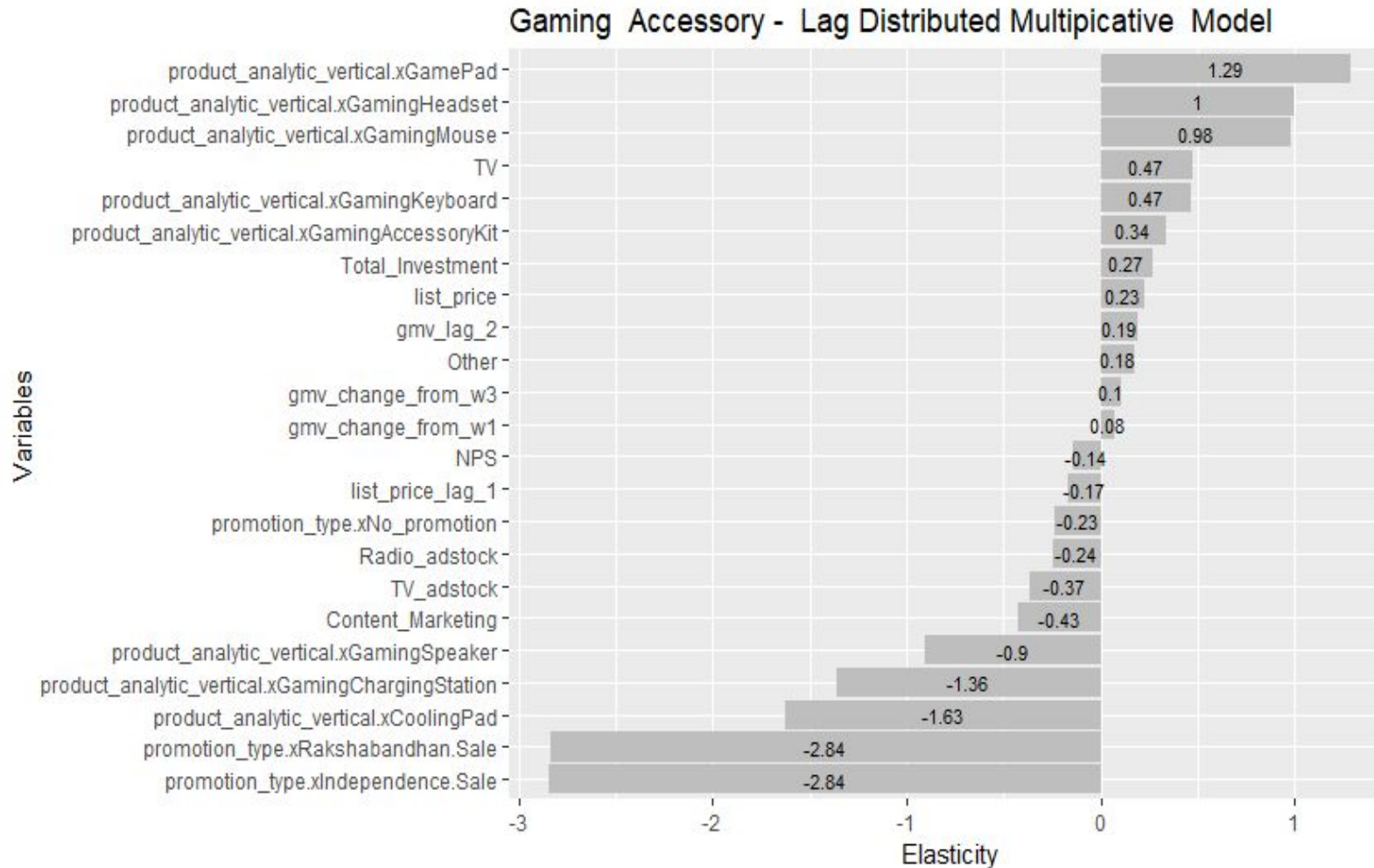
(4) Advertising adstock describes the prolonged or lagged effect of advertising on consumer purchase behavior. ... Adstock is a model of how response to advertising builds and decays in consumer markets.

GMV Elasticity across parameters for Camera Accessory



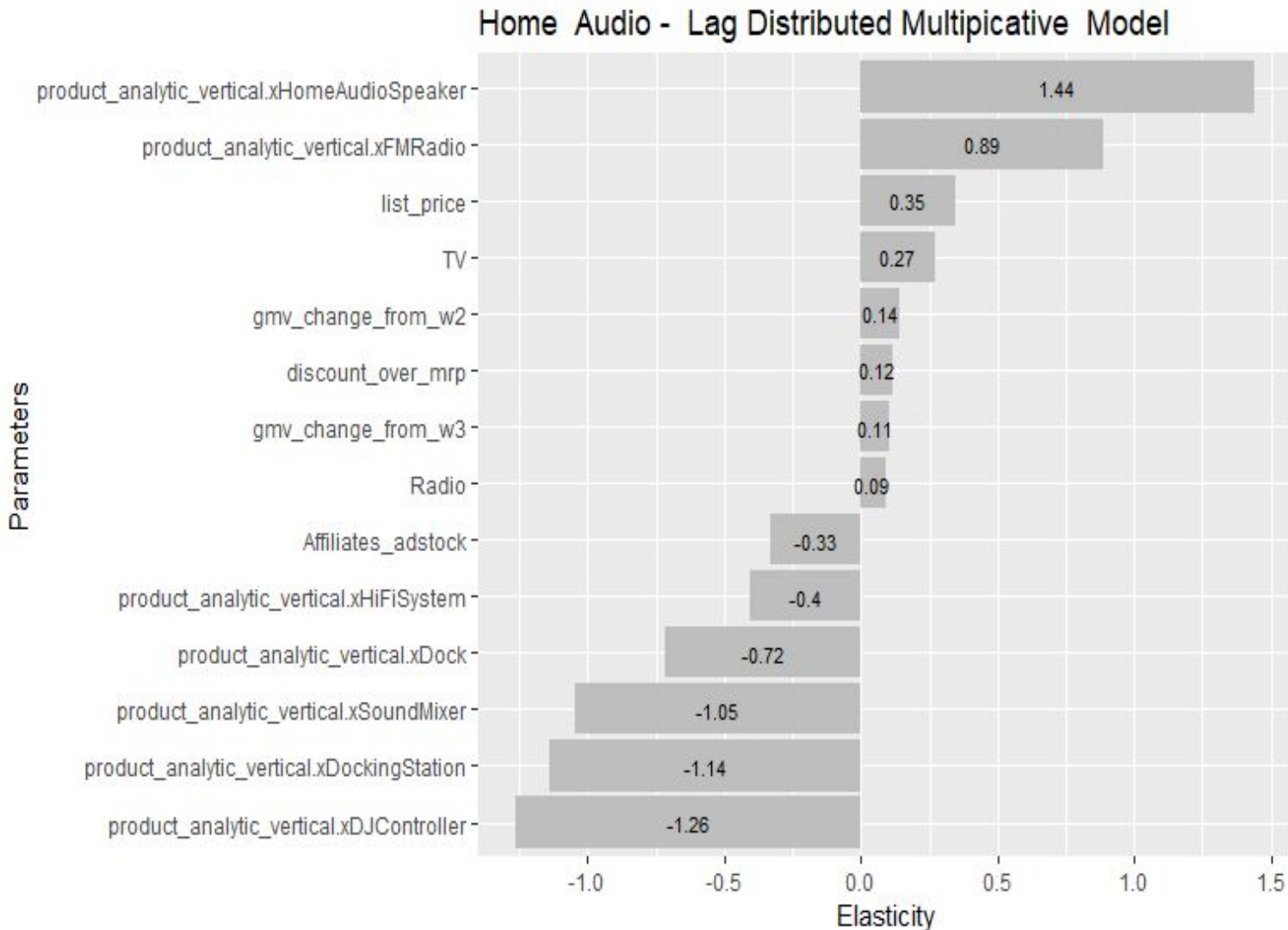
- Positive elasticity value implies with increasing value of the KPI, the GMV will increase.
- Positive elasticity value for Dussehra indicates that promotional efforts during Dussehra is likely to yield positive ROI.
- Elasticity of close to zero for Content Marketing and Sponsorship promo channels indicate relative inelasticity of Total GMV. Investment in these channels is likely to have very little impact on the Total GMV.
- Sale of Lens has a disproportionately large positive impact on total GMV, when compared to other camera SKUs. This is because average MRP of camera lens is higher than those of other camera accessory SKUs. Also the frequency of ordering of lens is higher than the other SKUs.

GMV Elasticity across parameters for Gaming Accessory



- Positive elasticity value implies with increasing value of the KPI, the GMV will increase.
- Positivity elasticity for TV promo channel indicate the effectiveness of TV in pushing larger sales volume and corresponding GMV.
- Elasticity of close to zero for gmv_change_from_w3 and gmv_change_from_w1 indicate that GMV doesn't change in a significant manner in a distributed lag manner post investment in a promo channel.
- Sales during Raksha Bandhan and Independence Day for gaming accessories has a negative impact on GMV. This can be due to high discount over MRP that brings down the overall GMV compared to other periods when discount percentages are lower. Hence excessive discounting should be avoided during these periods.

GMV Elasticity across parameters for Home Audio



- Positive elasticity value implies with increasing value of the KPI, the GMV will increase.
- Positivity elasticity for Home Audio Speaker and FM Radio indicates that the ecommerce recommendation engine should include these SKUs in promo newsletter and mobile app alerts.
- Elasticity of close to zero for gmv_change_from_w3 and gmv_change_from_w2 indicate that GMV doesn't change in a significant manner in a distributed lag manner post investment in a promo channel.
- Docking Station and DJ Controller are likely to be slow moving products contributing little to the Total GMV. To push more sales, discounts offered for these products inadvertently end up affecting Total GMV in the negative way. Hence discounts for these SKUs need to be reallocated.

Recommendations :

Camera Accessories

1. More investment in Camera Lenses, Battery and flash brings more revenue.
2. More discount over MRP , more returns.
3. Higher List price for products , more revenue.
4. Promote more around Independence day to fetch more sales for camera accessories.

Gaming Accessories

1. More investment in Game headset, Mouse, Gaming pad and keyboard result into higher returns.
2. More investment on TV as advertising channel brings more revenue.
3. Invest less in Radio, Tv and content marketing channels.
4. Try more promotion around Raksha Bandhan/Independence day for better festive sales

Home Audio Accessories

1. More investment in Home audio speaker, Fm radio result into higher returns.
 2. More investment on TV,Radio as advertising channel brings more revenue.
 3. More discount over MRP, more sales.
 4. Try more promotion during festive seasons.
- 