

ELECKART MARKET MIX MODELING

BY

SAAD SYED ADEEB (FACILITATOR)

ROHITASHWA GUHA

AGENDA

- BUSINESS OBJECTIVE AND DATA UNDERSTANDING
- DATA PREPARATION
- EDA
- KPI
- MODELS CREATED
- RECOMMENDATION

BUSINESS OBJECTIVE AND DATA UNDERSTANDING

•BACKGROUND- BUSINESS UNDERSTANDING

•ELECKART IS AN E-COMMERCE FIRM BASED OUT OF ONTARIO, CANADA SPECIALISING IN ELECTRONIC PRODUCTS. OVER THE LAST ONE YEAR, THEY HAD SPENT A SIGNIFICANT AMOUNT OF MONEY ON 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 THE 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

•OBJECTIVES

- TO CREATE MARKET MIX MODELS FOR THREE PRODUCT SUBCATEGORIES - CAMERA ACCESSORY, HOME AUDIO AND GAMING ACCESSORY.
- THE MODELS ARE TO BE BUILT ON WEEKLY BASIS FOR EACH OF THE ABOVE SUBCATEGORIES

BUSINESS OBJECTIVE AND DATA UNDERSTANDING

- **DATA UNDERSTANDING**
- THE FILES USED FOR THE ANALYSIS ARE:
 - CONSUMER ELECTRONICS –THE BASE DATA SET WHICH CONTAINS THE INFORMATION OF THE ORDERS PLACED.
 - PRODUCT LIST – THE DATASET THAT HAS THE FREQUENCY AND PERCENTAGE CONTRIBUTION OF EACH PRODUCT
 - MEDIA INVESTMENT – THE DATASET THAT GIVES INFORMATION ON THE DISTRIBUTION OF INVESTMENT SPENT FOR THE GIVEN TIME FRAME
 - MONTHLY NPS SCORE – THE DATASET THAT GIVES THE MONTHLY NPS SCORE AND STOCK INDEX FOR THE TIME FRAME.
 - ONTARIO-2015 AND ONTARIO-2016 – THE DATASETS THAT CONTAIN THE THE INFORMATION REGARDING THE CLIMATE FOR THE GIVEN TIME FRAME FOR THEIR RESPECTIVE YEARS

BUSINESS OBJECTIVE AND DATA UNDERSTANDING

- THE DATA CONSISTS OF THE FOLLOWING COLUMNS :
- FSN ID:** THE UNIQUE IDENTIFICATION OF EACH SKU
- ORDER DATE:** THE DATE ON WHICH THE ORDER WAS PLACED
- ORDER ID:** THE UNIQUE IDENTIFICATION NUMBER OF EACH ORDER
- ORDER ITEM ID:** SUPPOSE YOU ORDER 2 DIFFERENT PRODUCTS UNDER THE SAME ORDER , IT GENERATES 2 DIFFERENT ORDER ITEM IDs UNDER THE SAME ORDER ID; ORDERS ARE TRACKED BY THE ORDER ITEM ID.
- GMV:** GROSS MERCHANDISE VALUE OR REVENUE
- UNITS:** NUMBER OF UNITS OF THE SPECIFIC PRODUCT SOLD
- ORDER PAYMENT TYPE:** HOW THE ORDER WAS PAID – PREPAID OR CASH ON DELIVERY
- SLA:** NUMBER OF DAYS IT TYPICALLY TAKES TO DELIVER THE PRODUCT
- CUST ID:** UNIQUE IDENTIFICATION OF A CUSTOMER
- PRODUCT MRP:** MAXIMUM RETAIL PRICE OF THE PRODUCT
- PRODUCT PROCUREMENT SLA:** TIME TYPICALLY TAKEN TO PROCURE THE PRODUCT

DATA PREPARATION

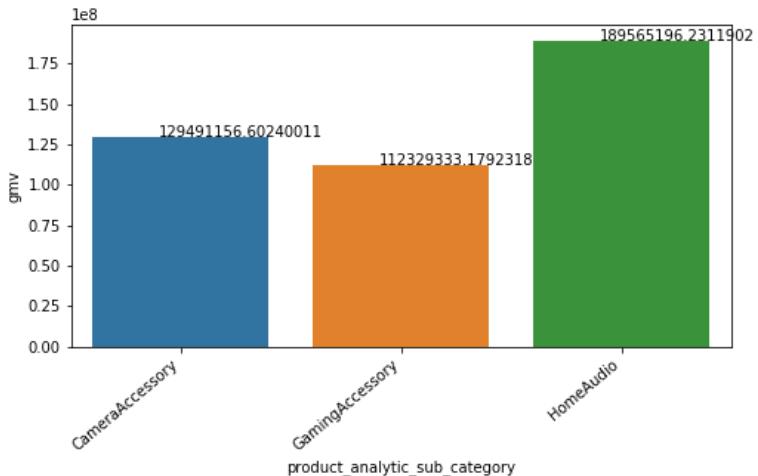
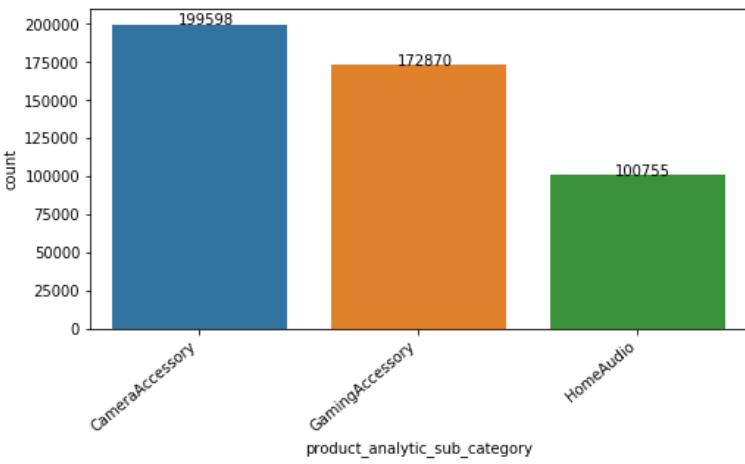
- BEFORE WE CREATE ANY MODELS ON THE DATA SET (CONSUMER ELECTRONICS IN THIS CASE), WE NEED TO FIRST PREPARE THE DATA BY CLEANING,IMPUTING ,CREATING NEW FEATURES ETC.
- THE APPROACH FOLLOWED BY US TO ACHIEVE THIS IS :
 - CHECK THE DATASET FOR ANY TYPE MISMATCH AND CORRECT THE SAME
 - DROP ANY DUPLICATE RECORD IF PRESENT .
 - REMOVE THE NULL VALUES BY EITHER DROPPING COLUMNS OR BY IMPUTING IT
 - REMOVING ROWS WITH PRODUCT MRP = 0
 - REMOVE ROWS WITH GMV = 0
 - VALUES ARE NEGATIVE FOR CUSTID,PINCODE,FSNID CORRECT THE VALUES FOR THE SAME.
 - ONLY TAKE RECORDS WHICH BELONG TO THE DESIRED TIME FRAME. (07/2015 TO 06/2016).
 - REMOVE ROWS WHICH HAVE GMV >MRP*UNITS AS IT'S NOT POSSIBLE

DATA PREPARATION

- CREATE NEW COLUMNS TO CHECK IF IT'S A PAY DAY OR A SPECIAL HOLIDAY
- CHECK FOR ANY OUTLIERS IN THE DATASET BASED ON THE FREQUENCY OF EACH VALUES AND REMOVE THE OUTLIERS
- CREATE NEW COLUMN TO STORE WEEKNUMBER WHICH WILL BE USED TO CONVERT FROM MONTHLY DATA TO WEEKLY DATA
- CAP THE PRODUCT PROCUREMENT SLA TO 30 DAYS
- CAP THE SLA TO 15 AND REMOVE THE NEGATIVE VALUES
- SEPARATE THE DATA INTO THREE DATA SETS ONE FOR EACH OF THE THREE PRODUCT SUB-CATEGORIES TO PERFORM ANALYSIS (CAMERA ACCESSORY, HOME AUDIO AND GAMING ACCESSORY).
- CONVERT THE DATA FROM MONTHLY TO WEEKLY DATA
- MERGE ALL THE THREE DATA FRAMES WITH THE OTHER FILES IN ORDER TO CREATE ONE FINAL DATA FRAME FOR EACH PRODUCT WHICH WILL BE USED FOR EDA AND USED TO CHOSE KPIs.

DATA PREPARATION

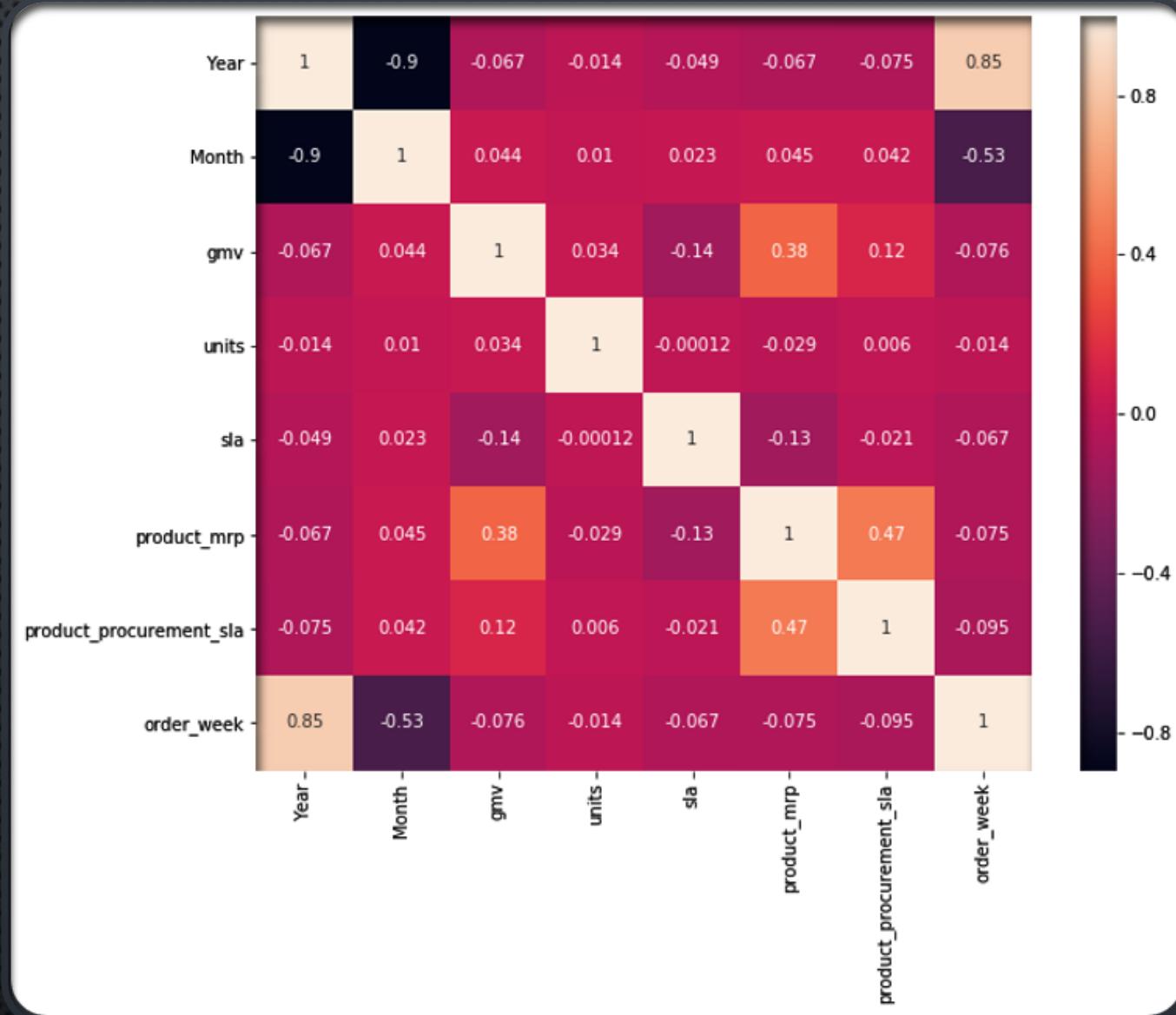
- OBSERVATION MADE DURING DATA PREPARATION:
- REMOVAL OF DUPLICATE RECORDS BASED ON THE COLUMN ORDER_ID.
- WHEN WE CALCULATE THE COLUMN WISE NULL VALUES, WE CAN CONCLUDE THAT 'DELIVERYBDAYS' AND 'DELIVERYCDAYS' HAVE 79% VALUES AS NULL VALUES ALSO HAS NEGATIVE VALUES THUS WE CAN DROP THESE COLUMNS.
- NOW, WE ANALYSE THE DATASET FOR OUTLIERS BY CHECKING THE COLUMNS SLA AND PRODUCT PROCUREMENT SLA. THERE IS NO DEFINITIVE VALUE WHICH CAN BE USED TO CAP THE NUMBER OF DAYS IT TAKES TO DELIVER THE PACKAGE. IN OUR MODEL, WE HAVE KEPT THE VALUE AT 15.
- AFTER THIS, WE DIVIDE THE DATASET INTO THREE DATAFRAMES (ONE FOR EACH PRODUCT SUBCATEGORY THAT WE NEED TO ANALYSE)



EDA

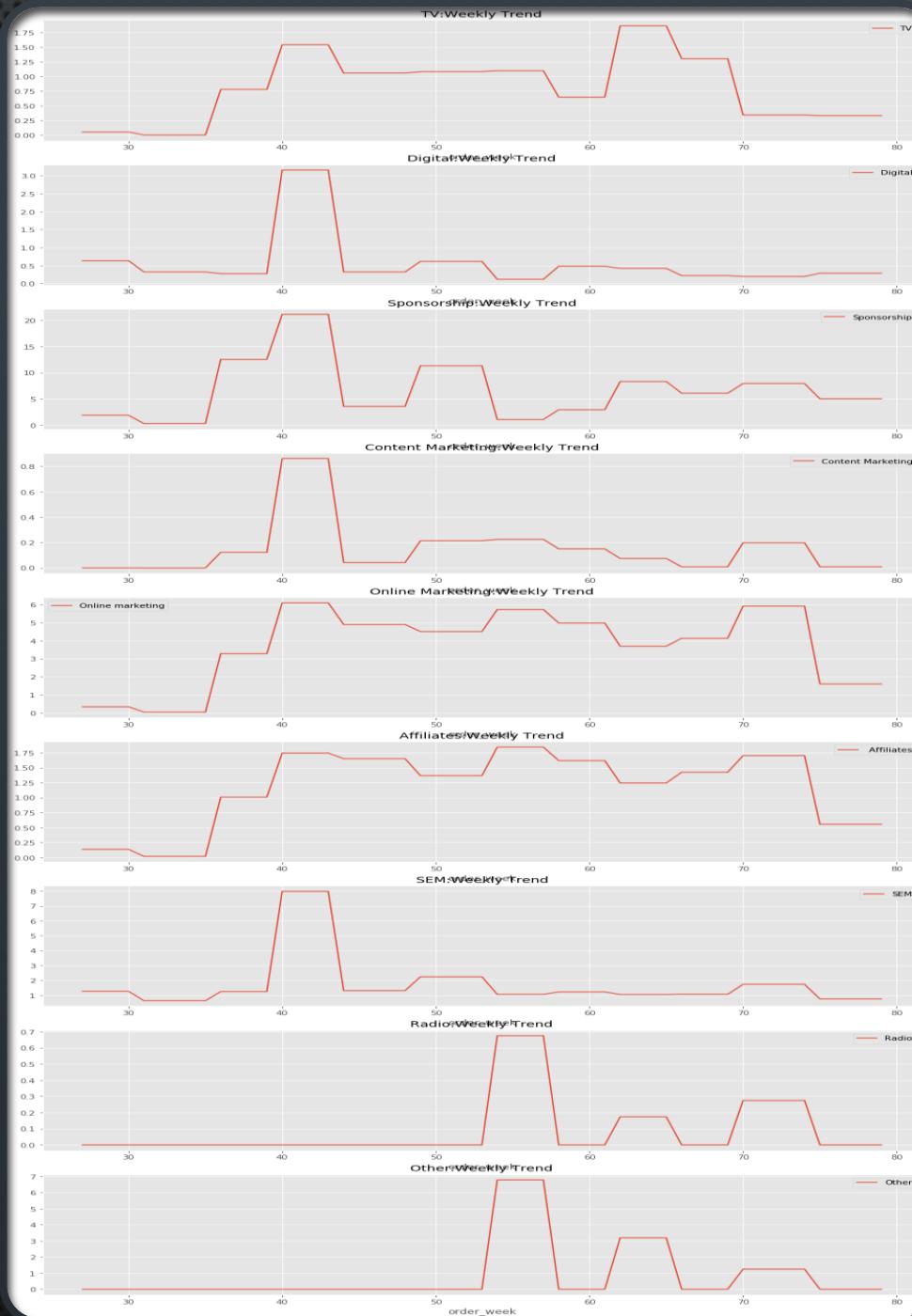
CAMERA ACCESSORY DATASET HAS THE HIGHEST NUMBER OF ORDERS WHILE
THE HOME AUDIOS DATASET HAS THE HIGHEST REVENUE GENERATED

EDA



THERE IS A HIGH CORRELATION
BETWEEN MONTH & YEAR, MONTH &
ORDER_WEEK AND MODERATE
CORRELATION BETWEEN PRODUCT_MRP
& GMV AND PRODUCT_MRP &
PRODUCT PROCUREMENT_SLA.

EDA

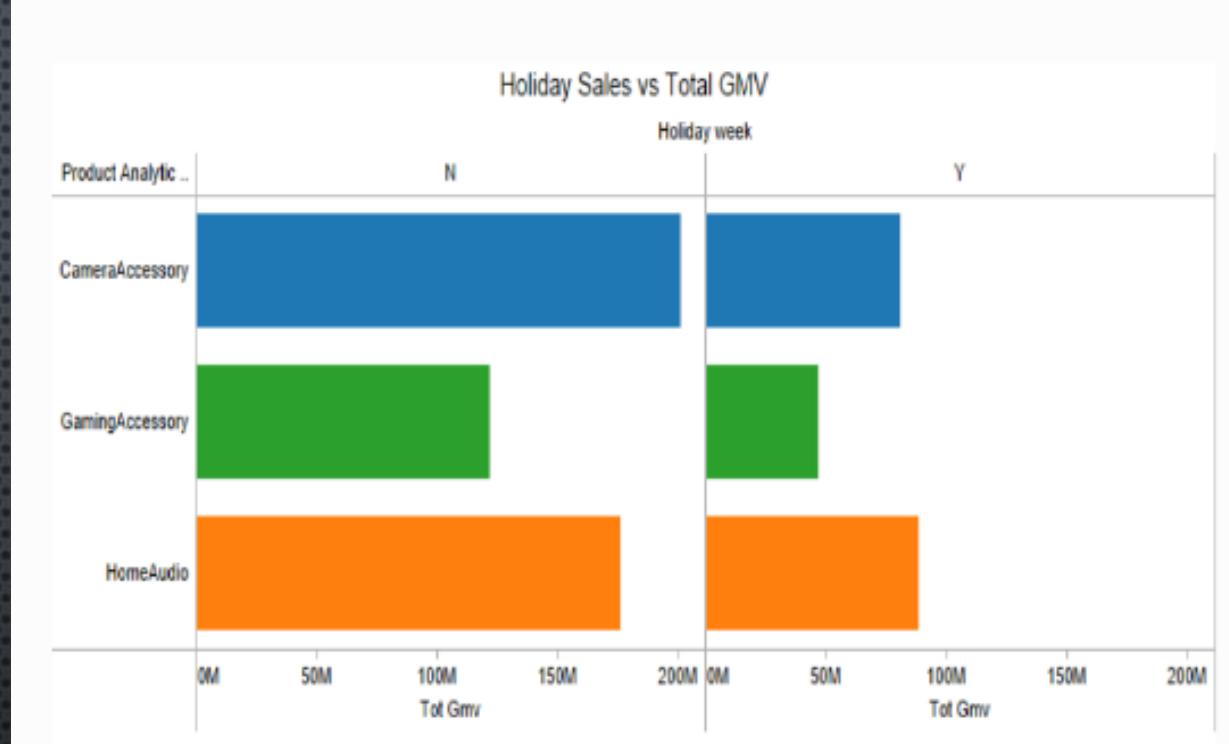


WE CAN CONCLUDE FROM THE FOLLOWING PLOT THAT:

- THE MAXIMUM NUMBER OF INVESTMENTS WERE DONE THROUGH SPONSORSHIP.
- MAXIMUM INVESTMENTS WERE DONE DURING THE DURATION OF WEEKS 39-45.

EDA

- DURING HOLIDAYS OR SPECIAL DAYS THE SALES INCREASE AND THUS THEY PROVE TO BE THE MOST PROFITABLE.
- CAMERA ACCESSORY GENERATES THE MAXIMUM REVENUE



EDA



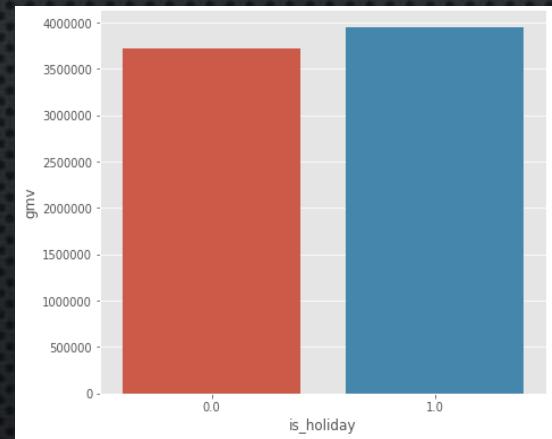
HIGHEST NUMBER OF ORDERS AS WELL AS THE HIGHEST REVENUE IS OBSERVED IN WEEK 42.

HIGHEST COUNT IS FOR HOME AUDIO WHILE REVENUE IS FOR GAMING ACCESSORY.

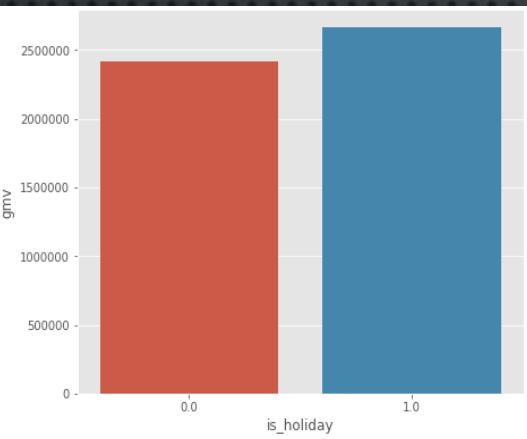
EDA

GRAPH SHOWING REVENUE SUM FOR BOTH HOLIDAYS AS WELL AS NON HOLIDAYS

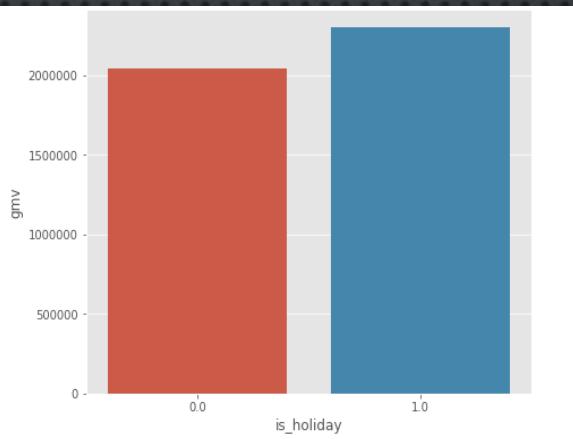
Home



Camera

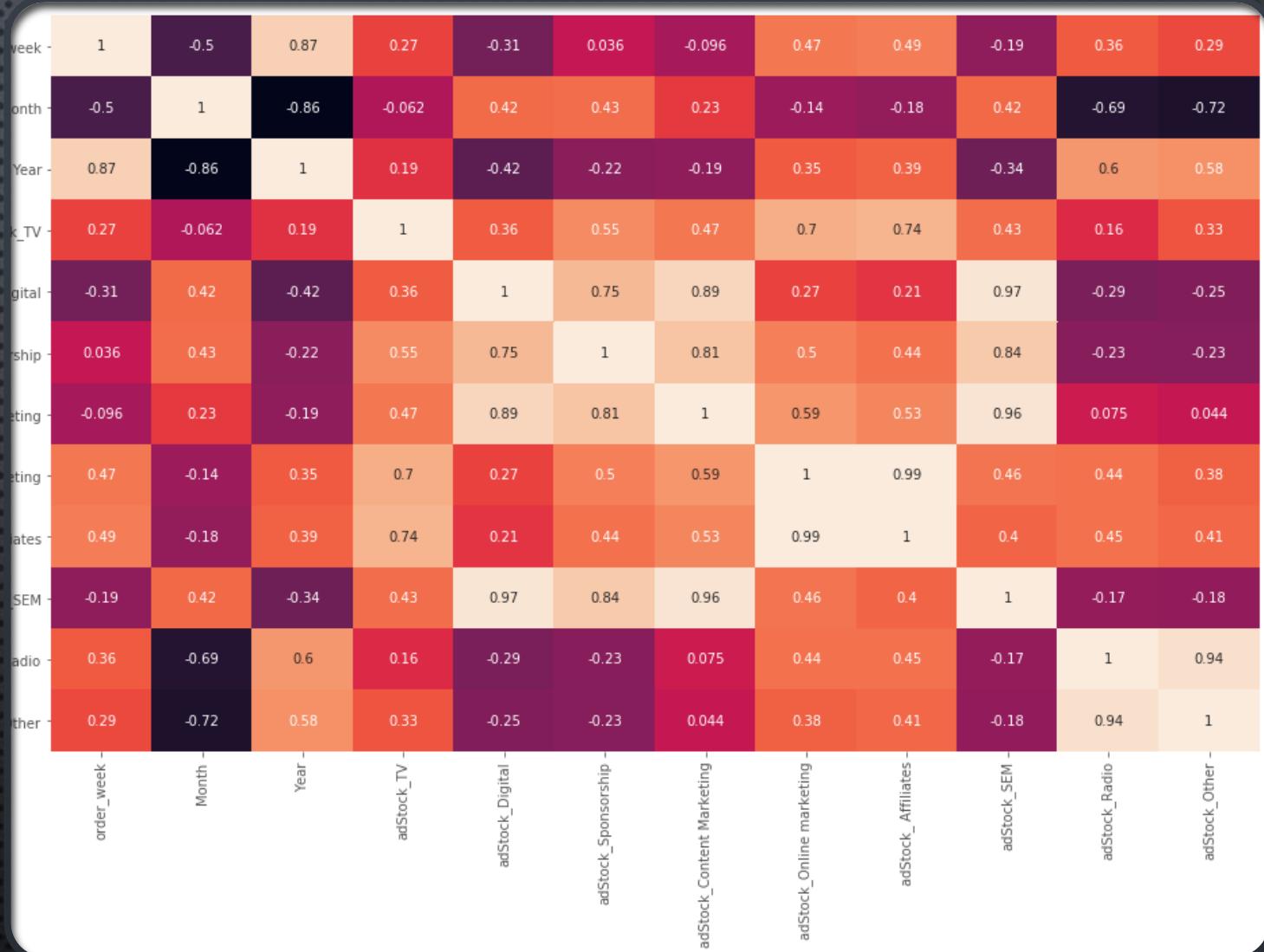


Gaming

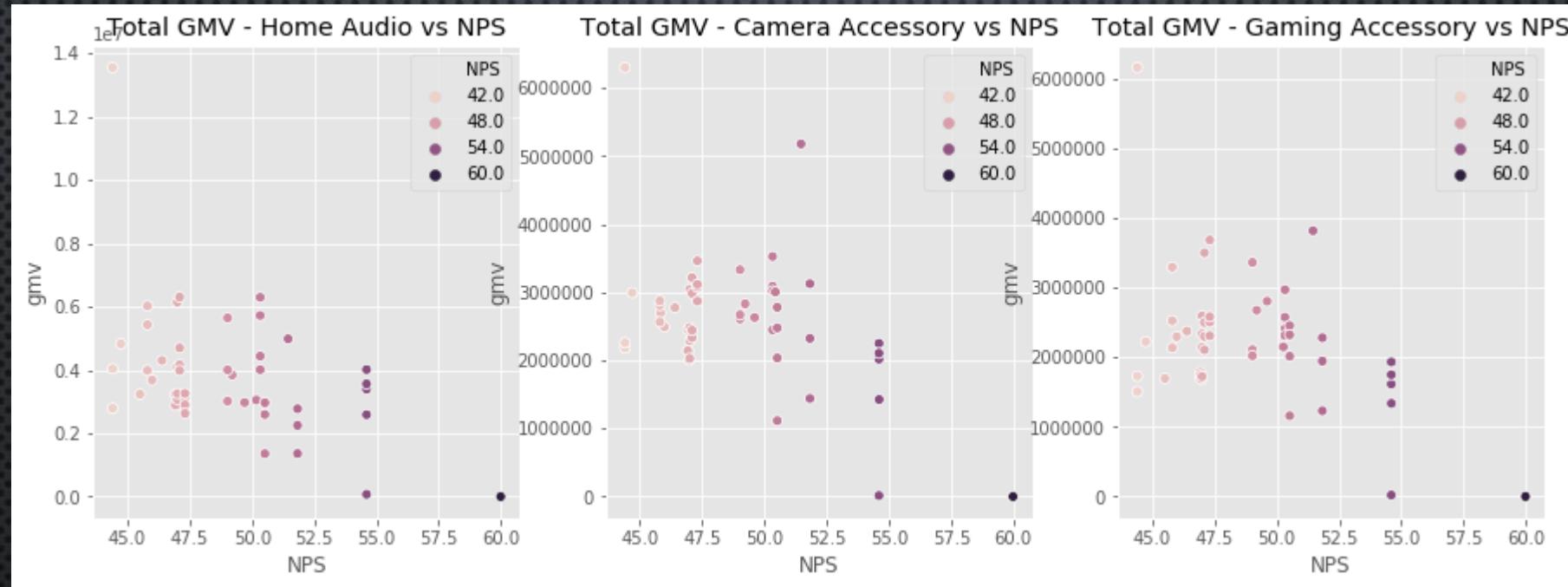


EDA

- AS WE CAN SEE, THERE IS A HIGH CORRELATION BETWEEN THE FOLLOWING:
- ORDER_WEEK & MONTH, YEAR & MONTH, ORDER_WEEK & YEAR
- ADSTOCK_AFFILIATES & ADSTOCK_ONLINE MARKETING
- ADSTOCK_OTHER & ADSTOCK_RADIO
- ADSTOCK_DIGITAL & ADSTOCK_SEM
- ADSTOCK_CONTENT MARKETING & ADSTOCK_SEM
- MONTH & ADSTOCK_RADIO, MONTH & ADSTOCK_OTHERS
- ADSTOCK_TV & ADSTOCK_AFFILIATES
- ADSTOCK_SPONSORSHIP & ADSTOCK_CONTENT MARKETING



EDA



NPS AND TOTAL GMV FOR ALL THE DIFFERENT SUB- CATEGORIES HAVE A NEGATIVE CORRELATION I.E. WITH INCREASE IN NPS, THE VALUE OF GMV IS DECREASING.

KPIS CHOSEN

- TOTAL GMV
- PAYMENTMODE INDICATOR
- PRODUCT PREMIUMNESS
- HOLIDAY WEEK
- SPECIAL WEEK
- ADSTOCK AFFILIATES
- ADSTOCK RADIO
- ADSTOCK SEM
- ADSTOCK CONSUMER ELECTRONICS
- PAYMENT WEEK
- SELLING PRICE
- DISCOUNT

CAMERA ACCESSORY - MODELS

MODEL	VARIABLES	ACCURACY
Linear	adStock_Digital,adStock_Online marketing,adStock_Affiliates,Binoculars, CameraBattery,Filter,Strap	56.70
Multiplicative	discount,total_premiumness, CameraTripod	97.5
Kyock	adStock_Radio,Binoculars, CameraBattery, CameraTripod,Filter,Strap	89.4
Distributed Lag Model	product_procurement_sla,discount,Binoculars,CameraTripod,Flash,Lens,Telescope,adStock_TV_lag1,adStock_Sponsorship_lag2,sla_lag3,product_procurement_sla_lag1	95.4

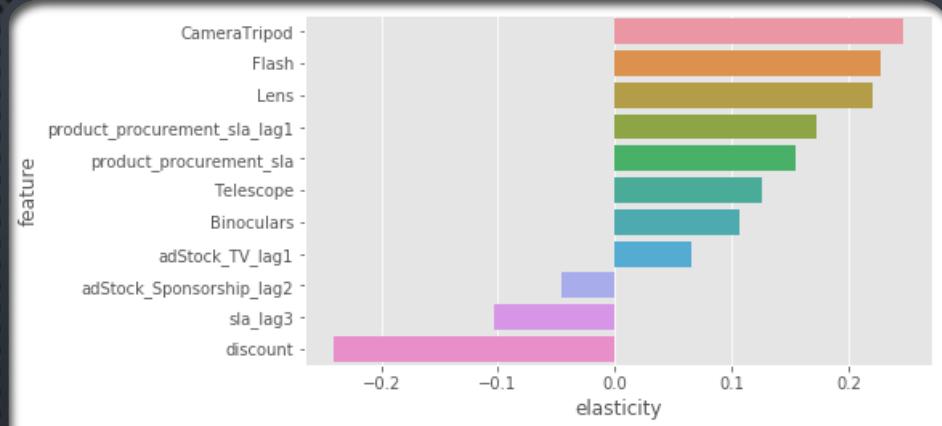
GAMING ACCESSORY-MODELS

MODEL	VARIABLES	ACCURACY
Linear	GamePad,GamingAccessoryKit,GamingHeadset,GamingMouse,GamingSpeaker	93.70
Multiplicative	GamingHeadset,GamingMouse,JoystickGamingWheel	79.00
Kyock	adStock_Radio,GamingAccessoryKit,GamingHeadset,GamingMouse,MotionController	91.1
Distributed Lag Model	GamePad,GamingHeadset,GamingKeyboard,GamingMemoryCard,GamingMouse,GamingSpeaker,gmv_lag3,product_procurement_sla_lag3,discount_lag2	93.40

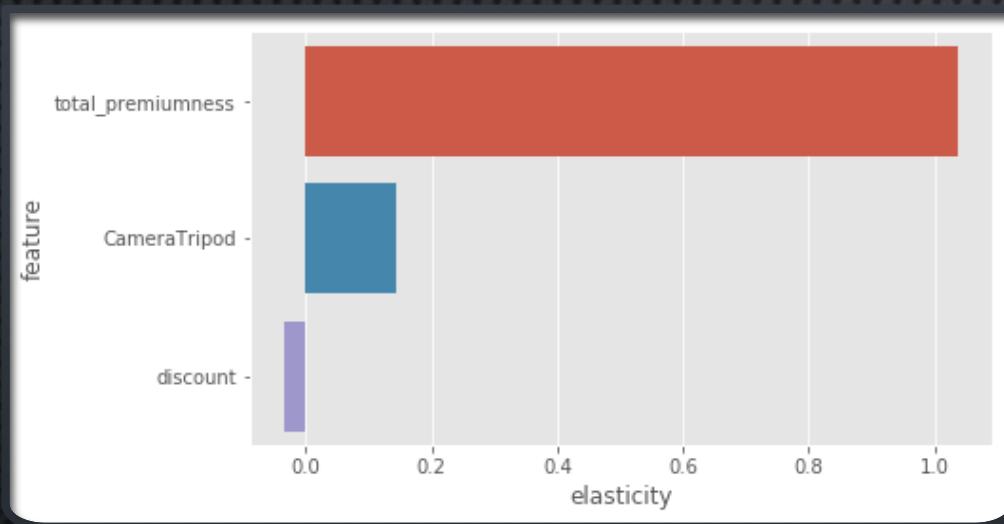
HOME AUDIO - MODELS

MODEL	VARIABLES	ACCURACY
Linear	BoomBox,HomeAudioSpeaker,VoiceRecorder	86.00
Multiplicative	adStock_Radio,Binoculars, CameraBattery, CameraTripod, Filter,Strap	89.00
Kyock	FMRadio,HomeAudioSpeaker,VoiceRecorder	98.30
Distributed Lag Model	FMRadio,HomeAudioSpeaker,adStock_Affiliates_lag3	97.80

CAMERA ACCESSORY – RECOMMENDATIONS



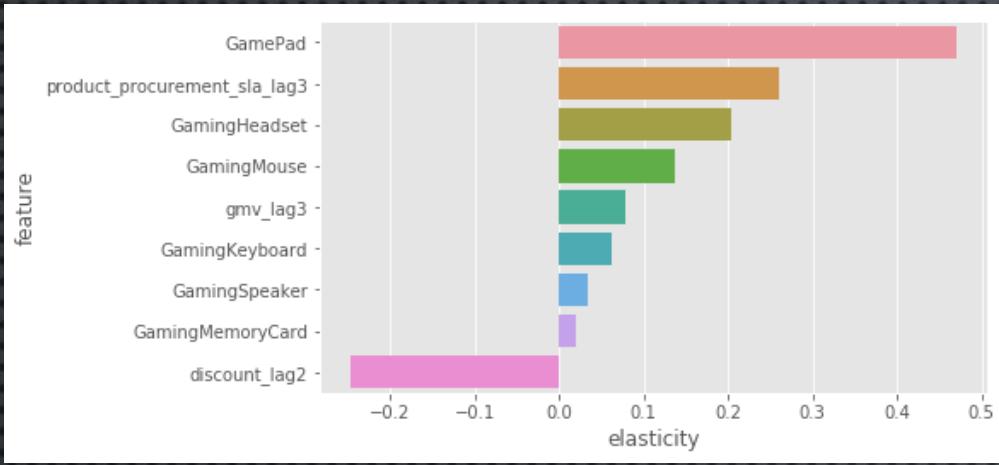
Elasticity of distributed lag model



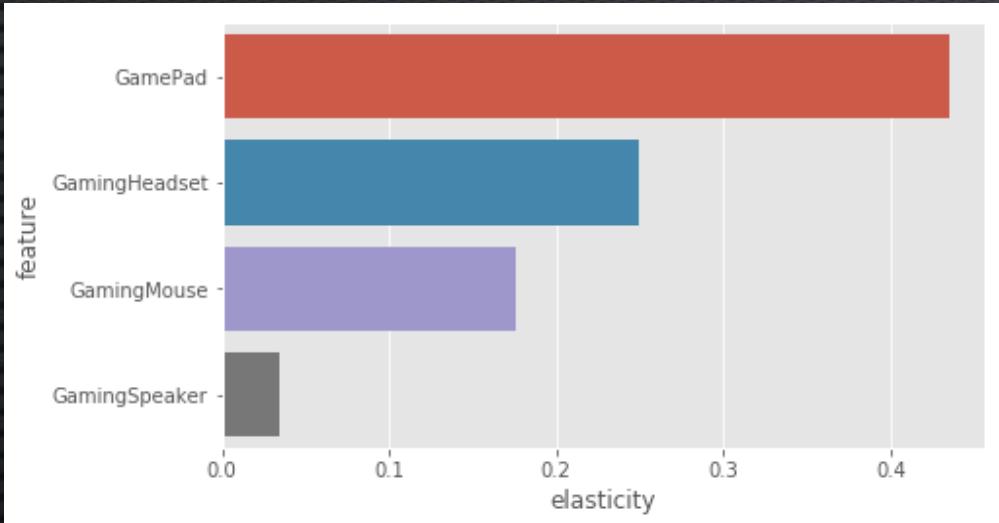
Elasticity of multiplicative model

- Multiplicative and distributed lag models are the 2 best models for camera accessory ,thus we'll be using the elasticity of the variables in this model to give the recommendations.
- Thus for Camera Accessory category it's a good idea to spend on marketing on TV channel and try and increase the sale of Camera tripod,Flash,Lens,Telescope,Binocular categories .
- Discounts should be reduced .
- More resources should be spent towards marketing of premium products.

GAMING ACCESSORY – RECOMMENDATIONS



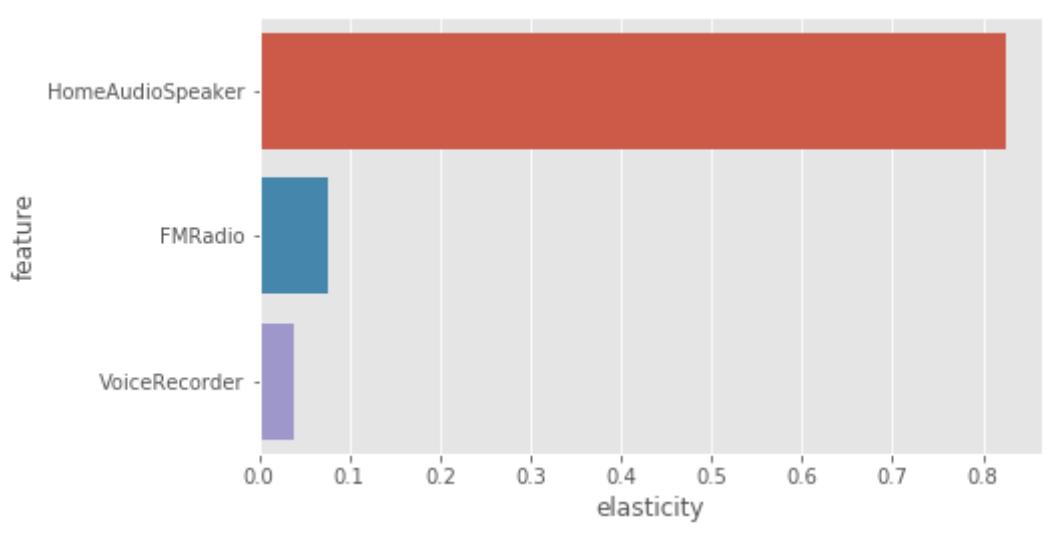
Elasticity of distributed lag model



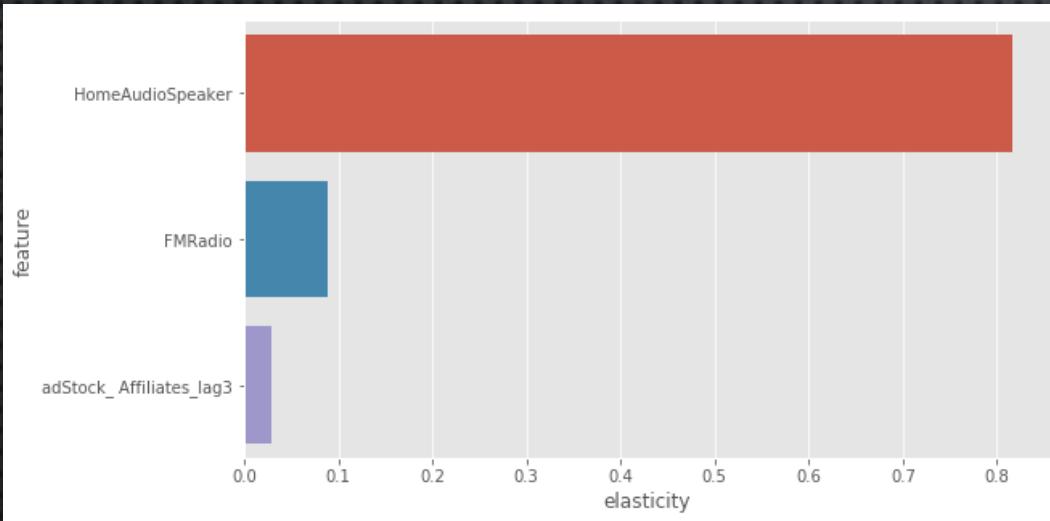
Elasticity of linear model

- Linear and distributed lag models are the 2 best models for gaming accessory ,thus we'll be using the elasticity of the variables in this model to give the recommendations.
- Thus for Gaming Accessory category it's a good idea to spend on marketing and try and increase the sale of Gaming headset,GamingMouse,GamePad.
- Discounts should be reduced from the last 2 weeks.

HOME AUDIO – RECOMMENDATIONS



Elasticity of Kyok model



Elasticity of distributed lag model

- Kyok and distributed lag models are the 2 best models for home audio ,thus we'll be using the elasticity of the variables in this model to give the recommendations.
- Thus for Home Audio category it's a good idea to spend on marketing and try and increase the sale of HomeAudioSpeaker ,BoomBox categories
- Resources should be spent on Affiliates as well marketing over the FM Radio channel.