ElecKart e-commerce Market Analysis

Nithya V

Suravaram Seshagiri Rao

Abstract

Business Understanding :

To develop a market mix model for ElecKart, to observe the actual impact of different marketing variables over the last year and to recommend the optimal budget allocation for different marketing levers for the next year.

Business Objectives :

To understand about ElecKart specialising in electronic products and its consumer order level data, media data, holiday data, stock index and monthly NPS scores.
Create market mix model based on the data and provide recommendations that help ElecKart in better budgeting.

Approach

1. Gather e-commerce domain knowledge and Market Mix Modelling (MMM)
2. Understand the consumer order level data, holiday, pay day, special sales, media, and NPS data.
3. Data Preparation : Clean, prepare and merge all the datasets into weekly level of 3 product sub categories – Camera Accessory, Gaming Accessory and Home Audio.
4. Feature Engineering: Listing price, Discount price, Lag variables, Ad Stock, Pay Day and holiday.
5. Exploratory Data Analysis
   - Univariates.
   - Bivariate.
   - Correlation.
   - Price Elasticity.
6. Market Mix Model Building
   - Baisc Linear Model
   - Koyck Model
   - Distributed Lag Model
   - Exponential attraction Model
   - Multiplicative Model
   - Hierarchical Model
7. Recommendation of the Model outcomes.

Data Understanding and Preparation

1. Importing ConsumerElectronics.csv into dataframe.

2. Finding the basic details like shape, columns and rows having Null values and duplicate values.

3. Rename the column names into a readable format.

4. Removing the outliers from the data frame.

5. Converting the columns into its corresponding data type.

6. Filter the data from July 2015 to June 2016.

7. Filter the data for three product subcategories - camera accessory, home audio and gaming accessory.

8. Create Week numbers from the order date in such a way that week number sequence is continued in January 2016.

9. Remove rows which have GMV as NaN.

10. Remove rows which have duplicate Order date, id, order item id, units and customer id.

11. Remove columns which have more null values and not helpful for the analysis.

12. Import Media data and other information.xlsx excel file into data frame. The Media and NPS data sheets should be imported to separate data frames.

13. In the media data frame remove the Radio and Other column which are having 75% NaN values.

14. Import the Canada, Ontario state Holiday calendar.

15. The climate data file ONTARIO-2015.csv and ONTARIO-2016.csv should be imported into separate data frames.

16. Finding the basic details like shape, columns and rows having Null values and duplicate values for the climate data.

17. Removing the columns that are having null values more than 70%.

18. Rename the column names into readable format and append the 2015 and 2016 data into single data frame.

Feature Engineering

1. From the MRP find price ending with 99 and mark it as Display_Price_ending_with_99 and others as Display_Price_others.

2. Create a Week data frame with the order date from consumer data and create the below KPI:
   1. Week, Month, Year, Day , TimeofDay, Daysinmonth.

3. Create SaleFlag to week data frame based on the sales date provided in media excel file.

4. Create WeekendFlag based on the day created in the above step.

5. Create PayDayFlag if the date is 1 or 15 in the month.

6. Create HolidayFlag if the date is in the holiday calendar imported.

7. Create listing price by GMV/units.

8. Create discount price by (MRP-listingprice)/MRP.

9. Create moving average, cumulative average and exponential average for GMV, listing price and discount price.

10. Create Lag variables for discount price and GMV.

11. Product premiums : Create 3 clusters each for Mass product, Premium product and Aspiring product based on product vertical, units, MRP and listing price.

12. Create dummies for the ProductVertical column.

13. Create a pivot table based using sum as aggregate function on GMV and other columns required and create index on ProductSubCategory and Week.

14. Create a pivot table based using mean as aggregate function on the ProductSubCategory, Week, SLA, MRP, ProcurementSLA using index on ProductSubCategory and Week and merge all the pivot table data frames created.

15. Create Ad stock variable for TV, Affiliates, Digital, Online and Content marketing.

16. Convert the monthly media data into daily and aggregate into a weekly level.

17. Convert the NPS into weekly and merge with media data.

18. Merge the climate data with the above data to create a final data set to create models and perform EDA.
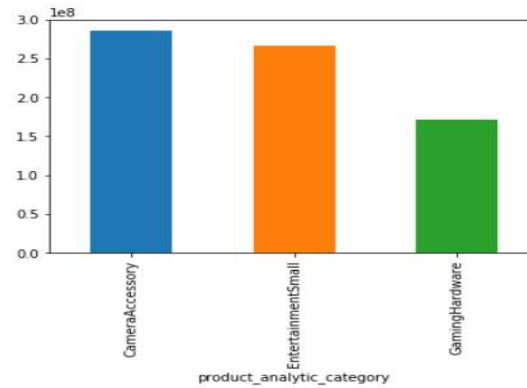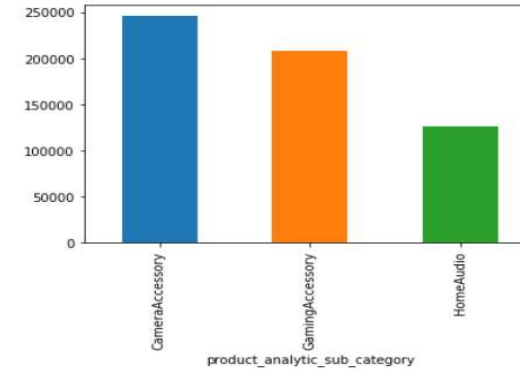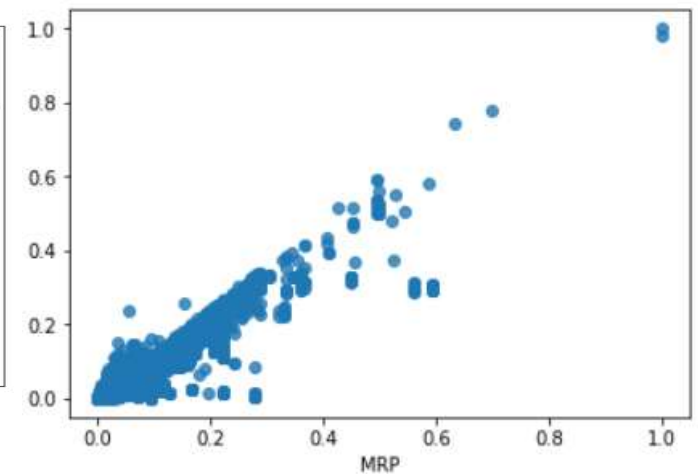
Exploratory Data Analysis – Univariate
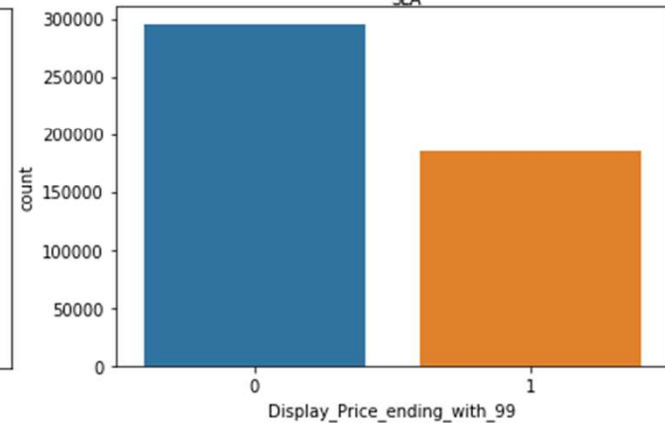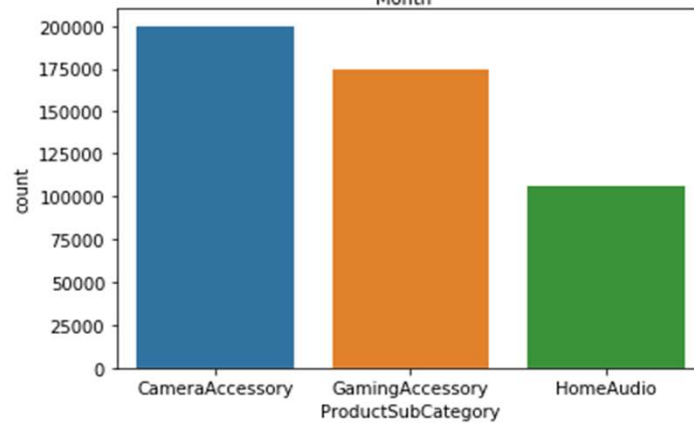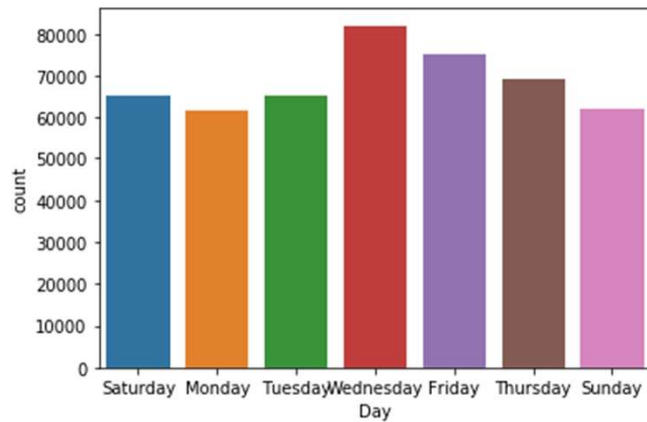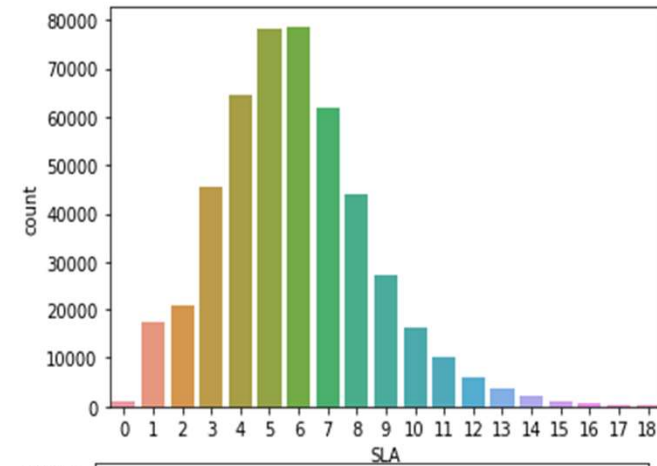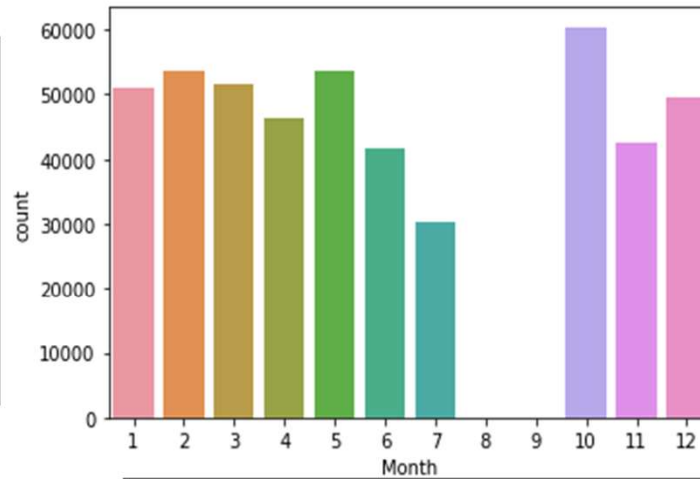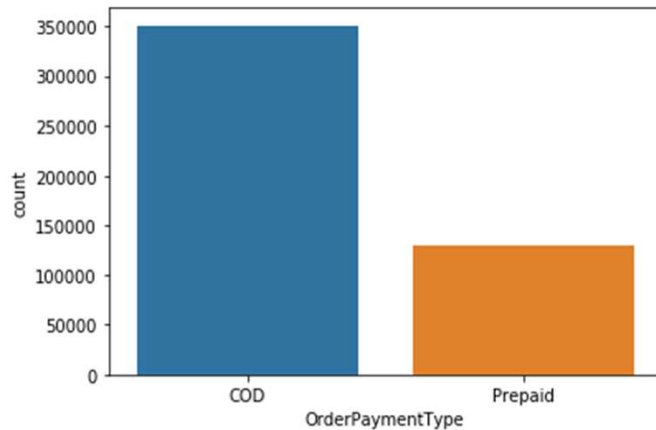
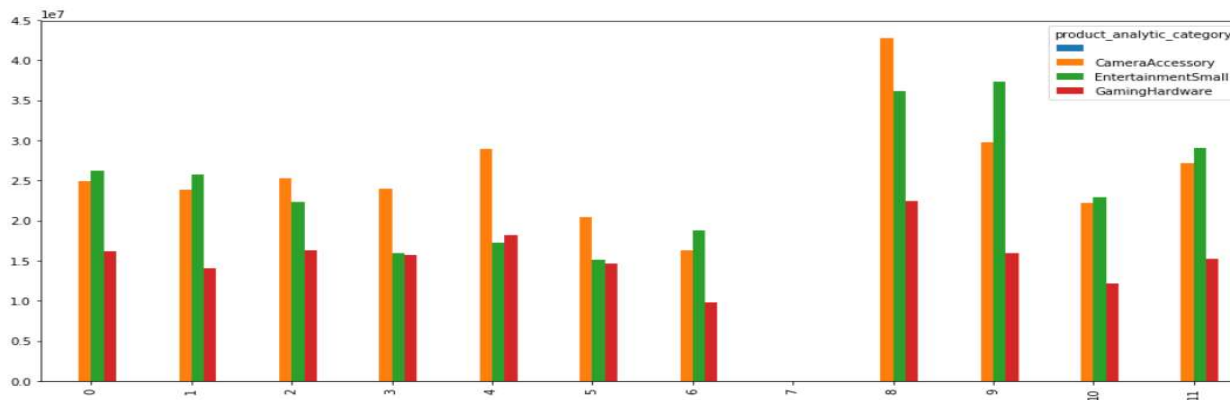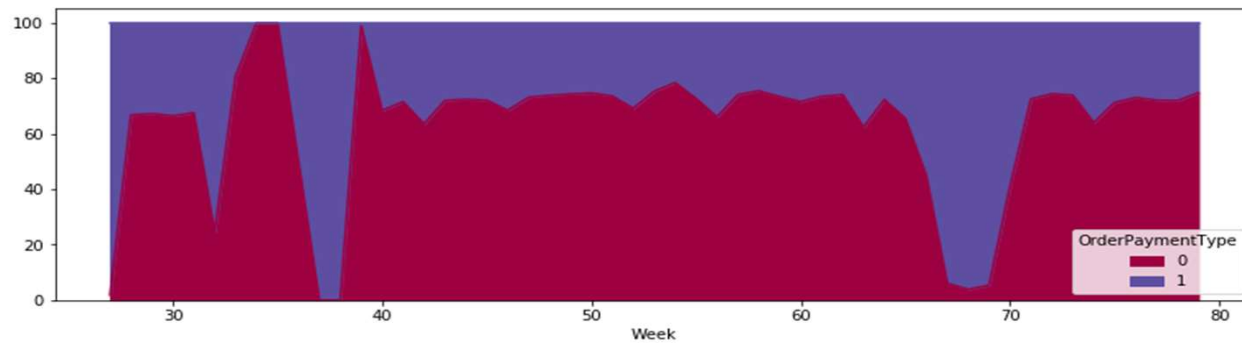# Exploratory Data Analysis – Bivariate 1
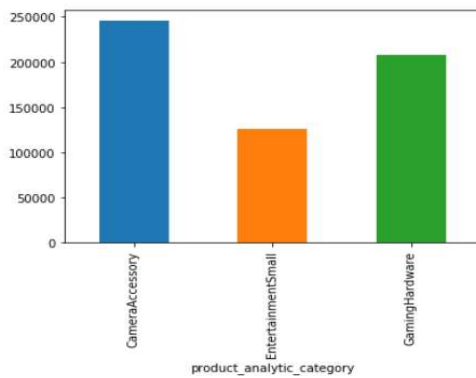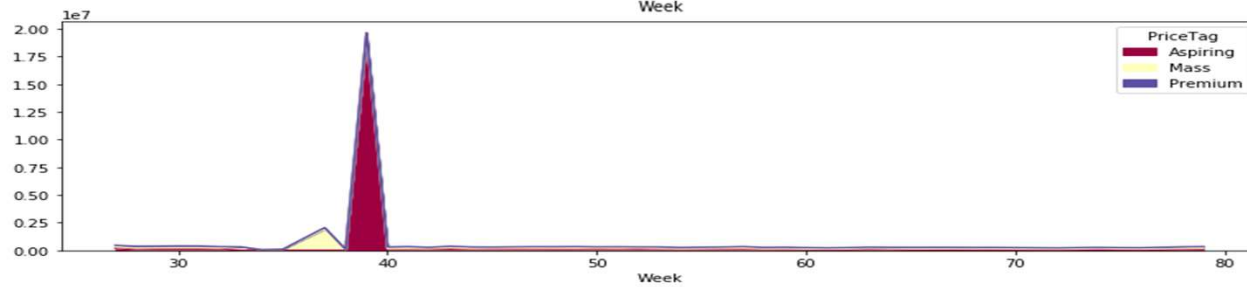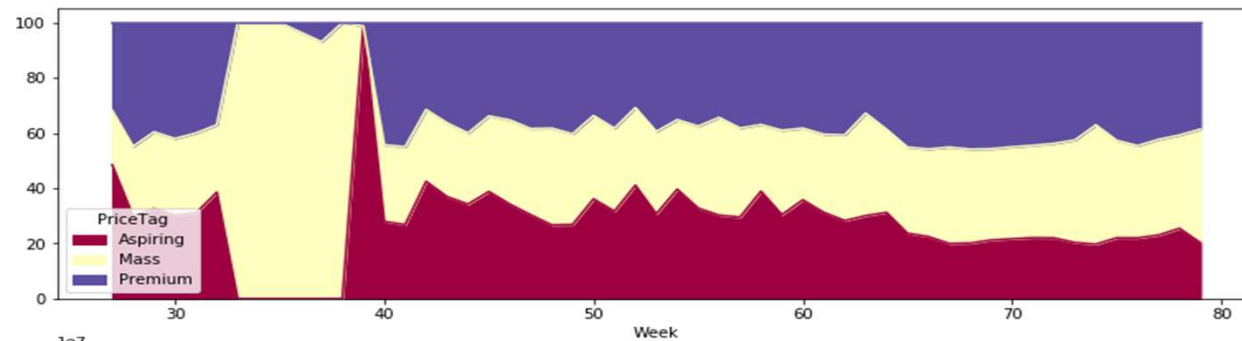

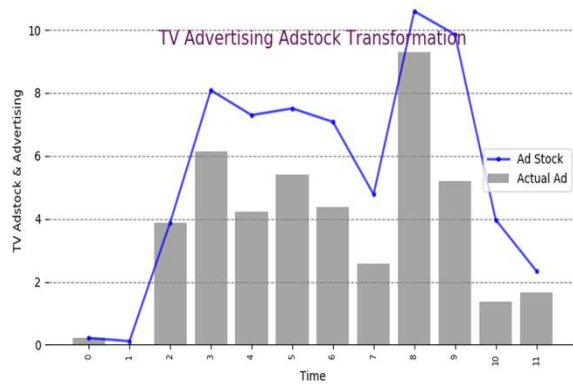
Product analytic category and units
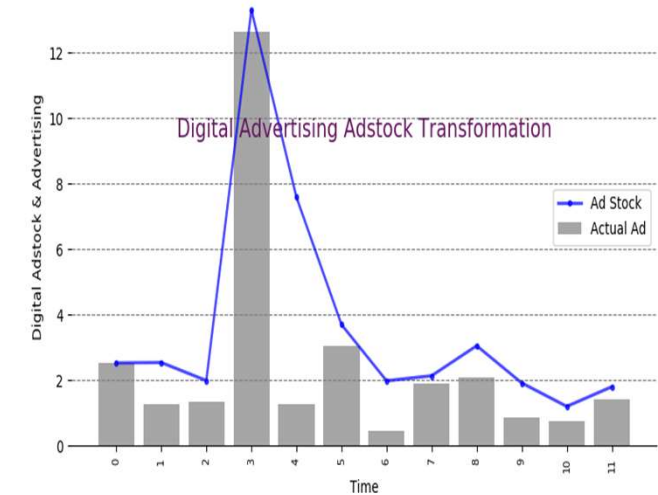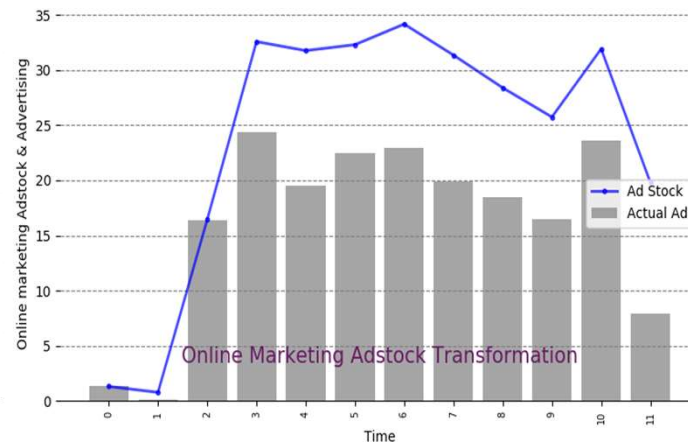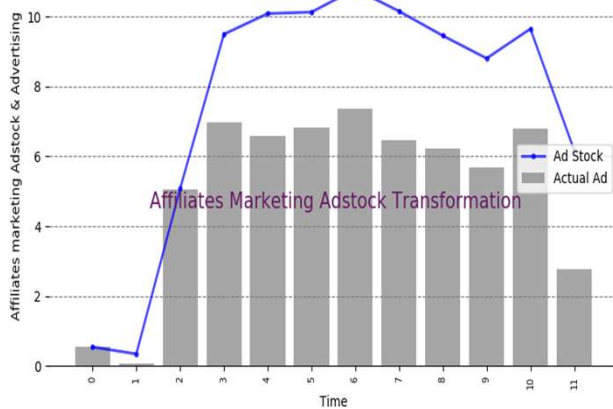

Product analytic category and GMV

Product analytic sub category and units
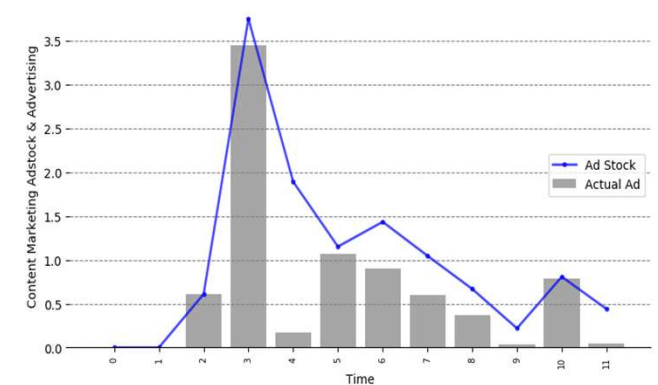
Exploratory Data Analysis – Bivariate 2

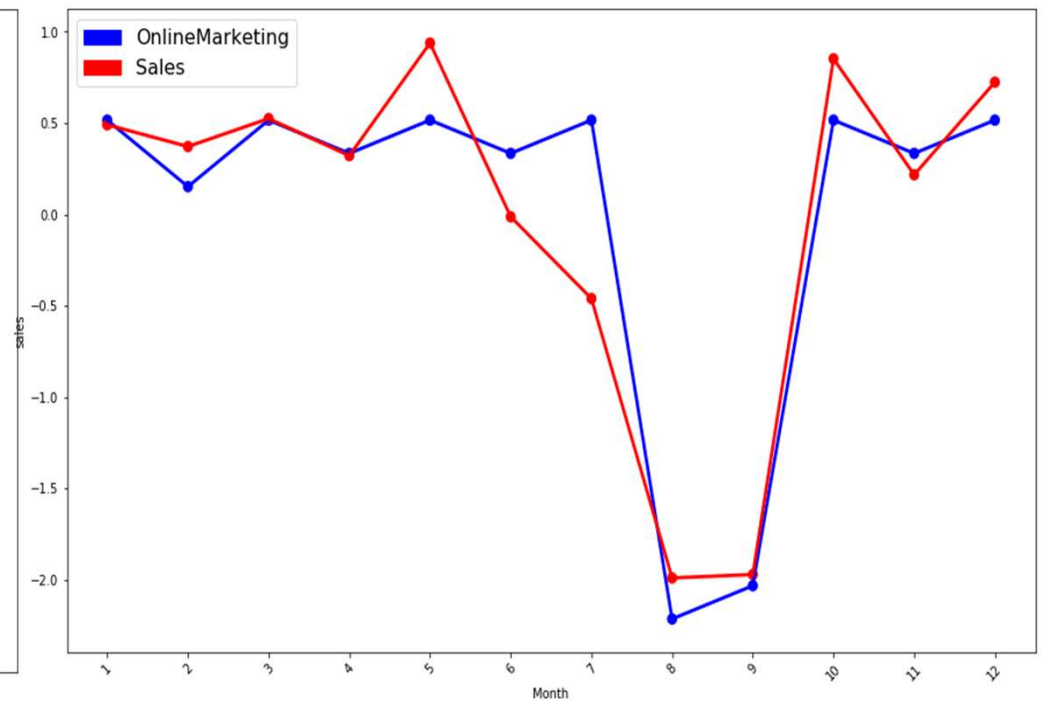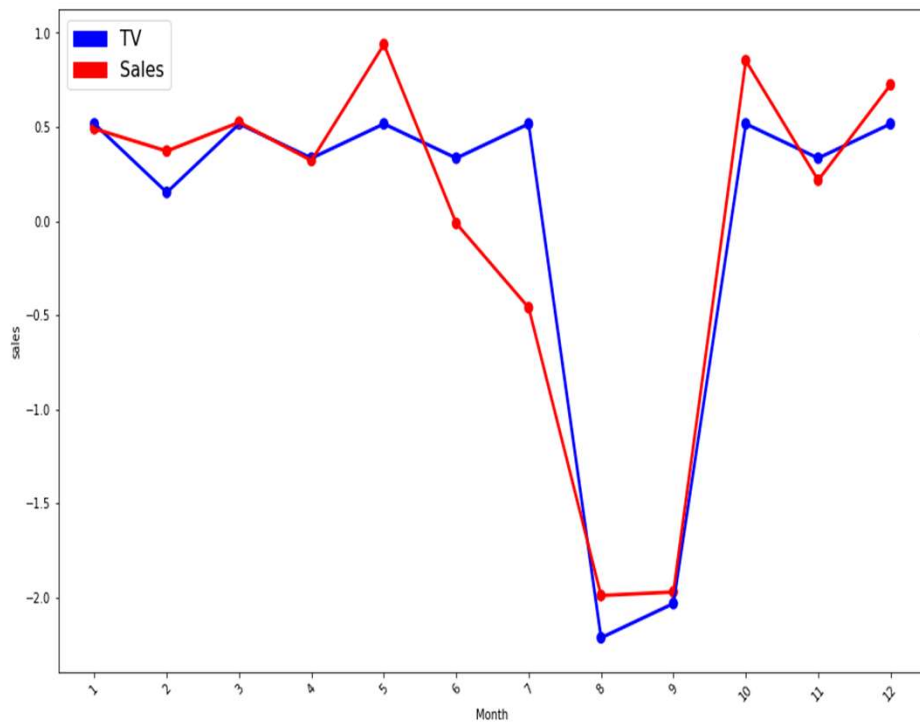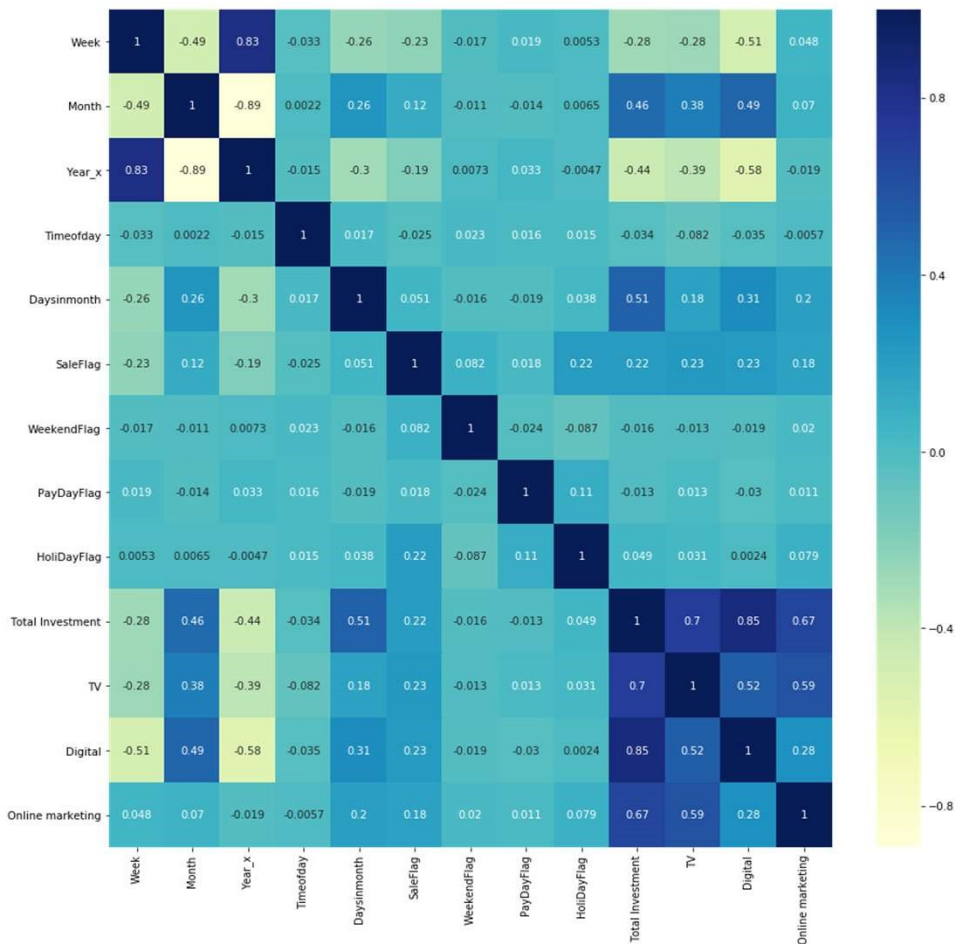# Exploratory Data Analysis – Ad Stock

Adstock variables are
created and ploted for
TV advertisements,
Digital,
Content marketing,
Online marketing and
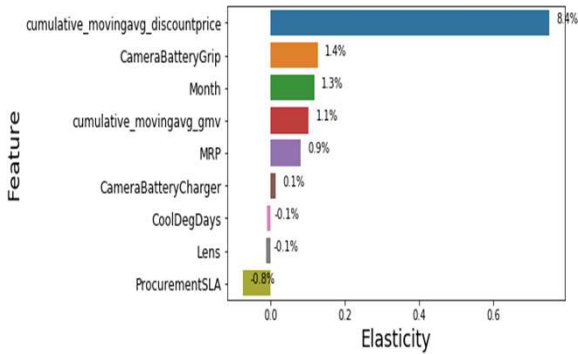Affiliates marketing

Exploratory Data Analysis - Correlation 2

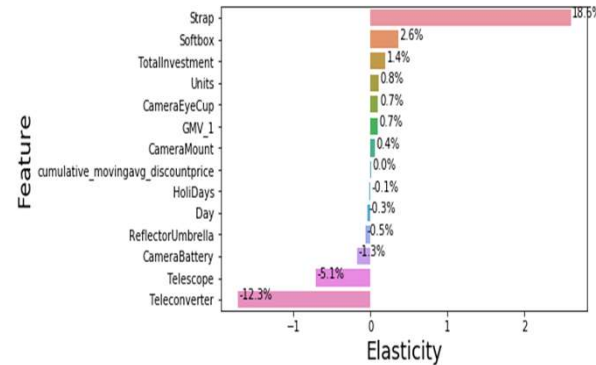Heat map correlation of Media and Holiday, Payday, Special Sales Day.

Variables which are highly correlated, having correlation greater than 0.90 are removed.
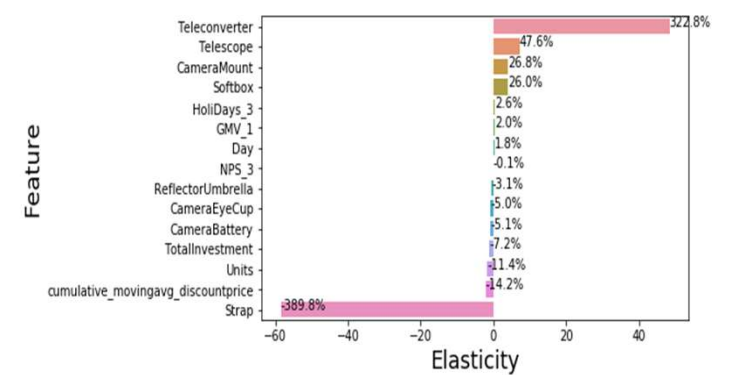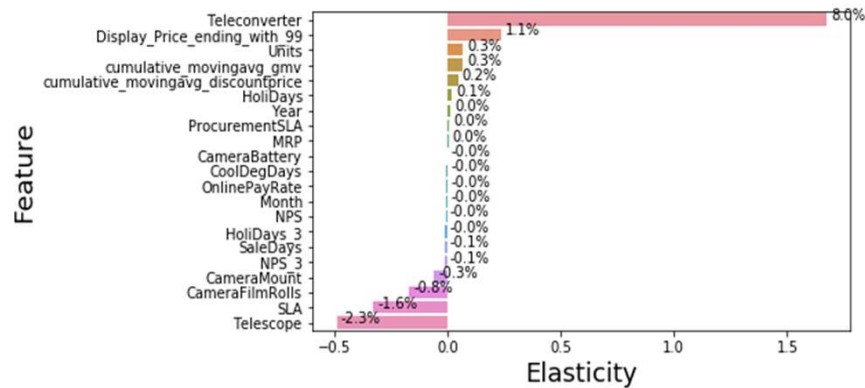
Camera Accessory Modelling

Linear Model Features
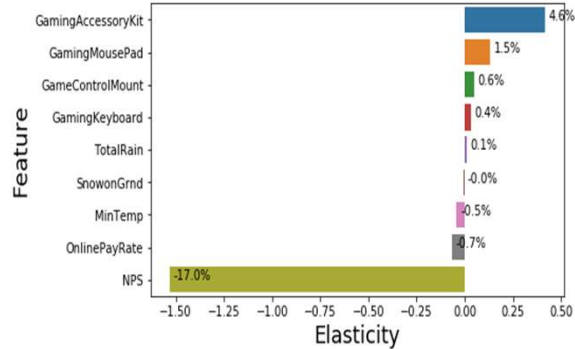

Kyock Model Features

Distributed Lag Model Features

Exponential/Multiplicative Model Features

Model	Camera - Key Features	R2 Test	R2 Train
Basic Linear	cumulative_movingavg_discountprice, cumulative_movingavg_gmv,	0.91	0.93
Koyck	Strap, Softbox,TotalInvestment, Units, CameraEyeCup,	0.98	0.98
Distributed Lag	Teleconverter,Telescope,TotalInvestment,CameraBattery,CameraEyeCup,CameraMount	0.98	0.98
Exponential/Multiplicative	Strap,ProcurementSLA,OnlinePayRate,cumulative_movingavg_gmv,CameraMount,Display_Price_ending_with_99,	0.99	1
Hierarchical	Units, cumulative_movingavg_discountprice, CameraMount, Strap, Teleconverter, OnlineMarketing, TV,GMV_1	0.89	0.99

Game Accessory Modelling

Linear Model Features

Kyock Model Features
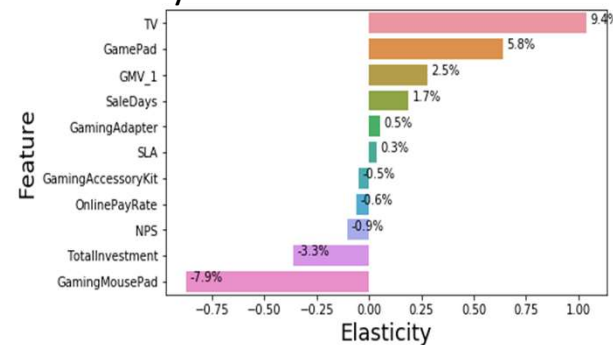
Distributed Lag Model Features
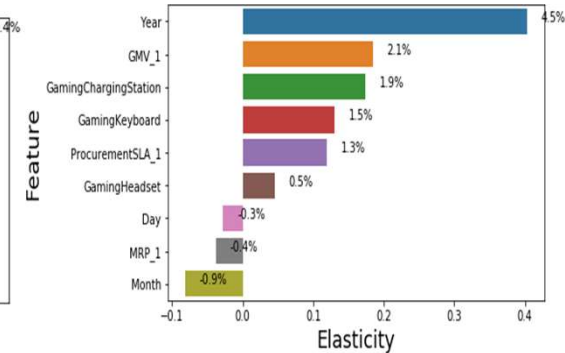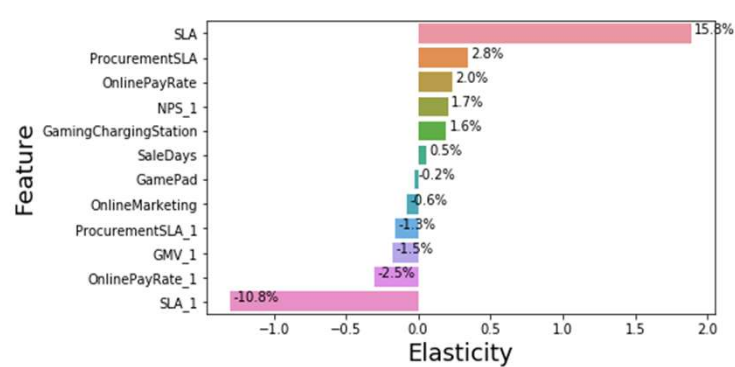

Exponential/Multiplicative Model Features

Model	Game Key Features	R2 Test	R2 Train
Basic Linear	GamingAccessoryKit,GamingMousePad,NPS,GameControlMount,GamingKeyboard	0.78	0.87
Koyck	TV,GamePad,GMV_1,SaleDays,GamingAdapter,GamingMousePad,GamingAccessoryKit,TotalInvestment	0.94	0.96
Distributed Lag	Year,GMV_1,GamingChargingStation,GamingHeadset,GamingKeyboard,ProcumentSLA_1	0.88	0.95
Exponential/Multiplicative	SLA,GamingChargingStation,GMV_1,HoliDays,NPS,GamePad,GamingMousePad ,TotalInvestment	0.95	0.99
Hierarchical	GamePad, GamingAccessoryKit, GamingAdapter,GamingChargingStation, GamingKeyboard,OnlineMarketing, TV,GMV_1	0.97	0.92
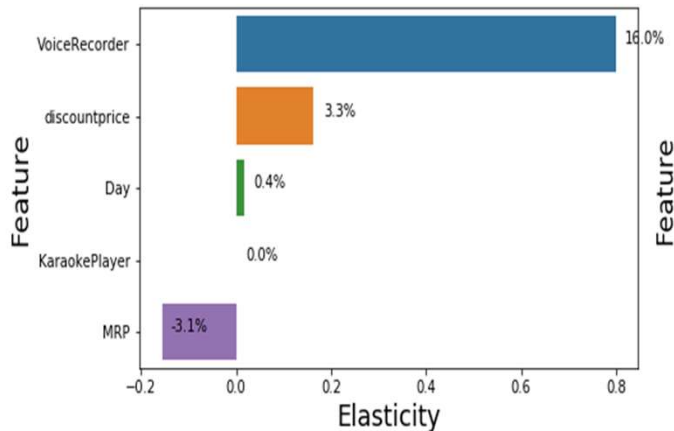
Home Audio Modelling

Linear Model Features

Kyock Model Features
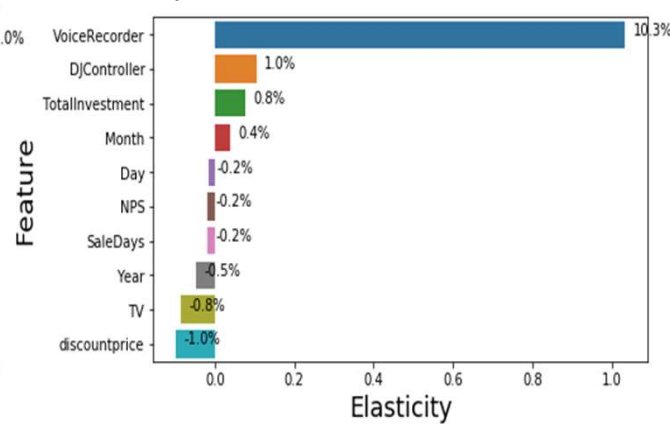
Distributed Lag Model Features
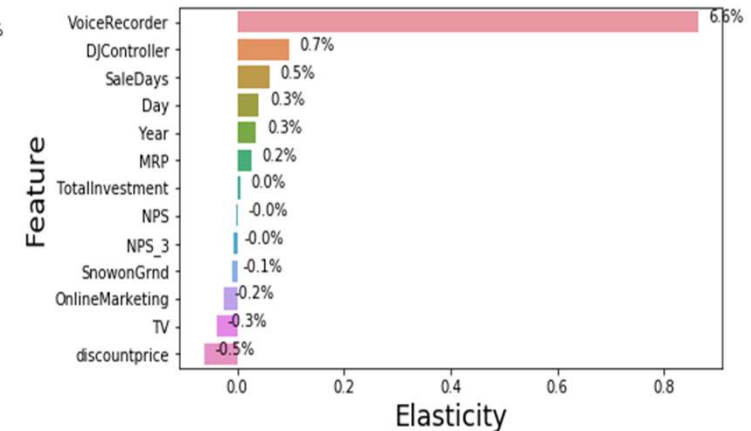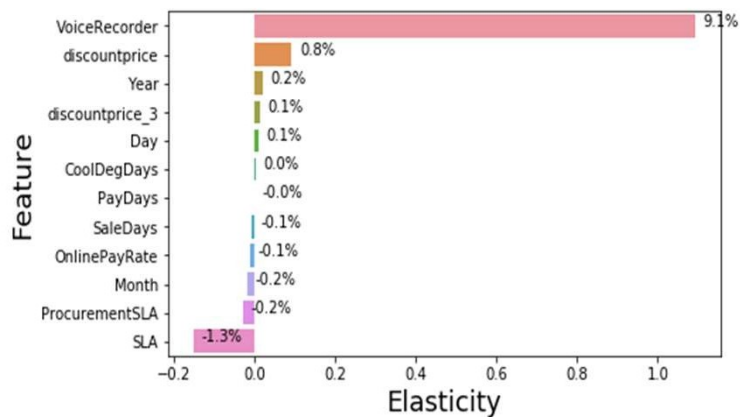

Exponential/Multiplicative Model Features

Model	HomeAudio Key Features	R2 Test	R2 Train
Basic Linear	VoiceRecorder,discountprice,MRP,Day and KaraokePlayer	0.98	0.98
Koyck	VoiceRecorder,DJController,TotalInvestment,Month,Day,NPS,SaleDays,TV,discountprice	0.99	0.99
Distributed Lag	VoiceRecorder,discountprice, TotalInvestment,DJController,SaleDays,TV,NPS,Day,Year,	0.99	0.95
Exponential/Multiplicative	VoiceRecorder,discountprice, Day,Year, CoolDegDays,PayDays,SaleDays,OnlinePayRate	0.99	0.99
Hierarchical	discountprice, DJController,VoiceRecorder, OnlineMarketing,SaleDays, TV,TotalInvestment	0.76	0.99

Recommendation

- If huge discounts then more Prepaid payment orders.

- Small discounts on aspiring products can result in a increase of sales drastically.

- Promotions on the products Strap, Teleconverter, Camera Mount, soft box will increase the sale value and advertising the same on TV will give better results.

- Advertising the products Game charging station, Gaming headset, Game pad and Gaming keyboard on TV will result in more sales.

- More discount on Voice recorder and DJ Controller will increase in the sale value.

- Display price ending with 99 for the camera products will increase the sales.