

IIT - BANGALORE

E-COMMERCE CAPSTONE

ELECKART

By:

Shubham Gupta, Ayush Garg and Sneha

OBJECTIVE



Company wants to relocate their budget optimally across different marketing levers



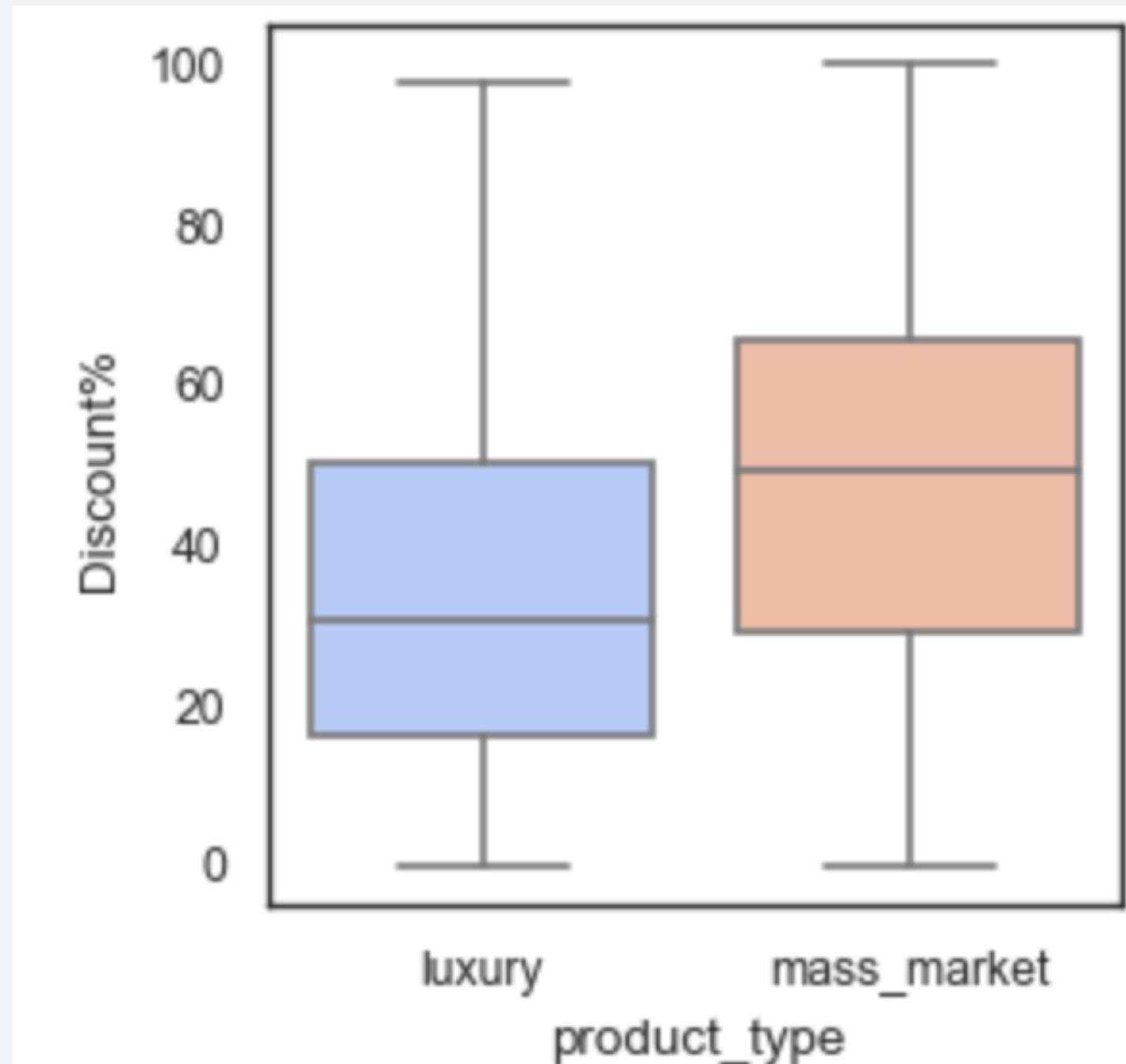
Build a market mix model



ANALYSIS

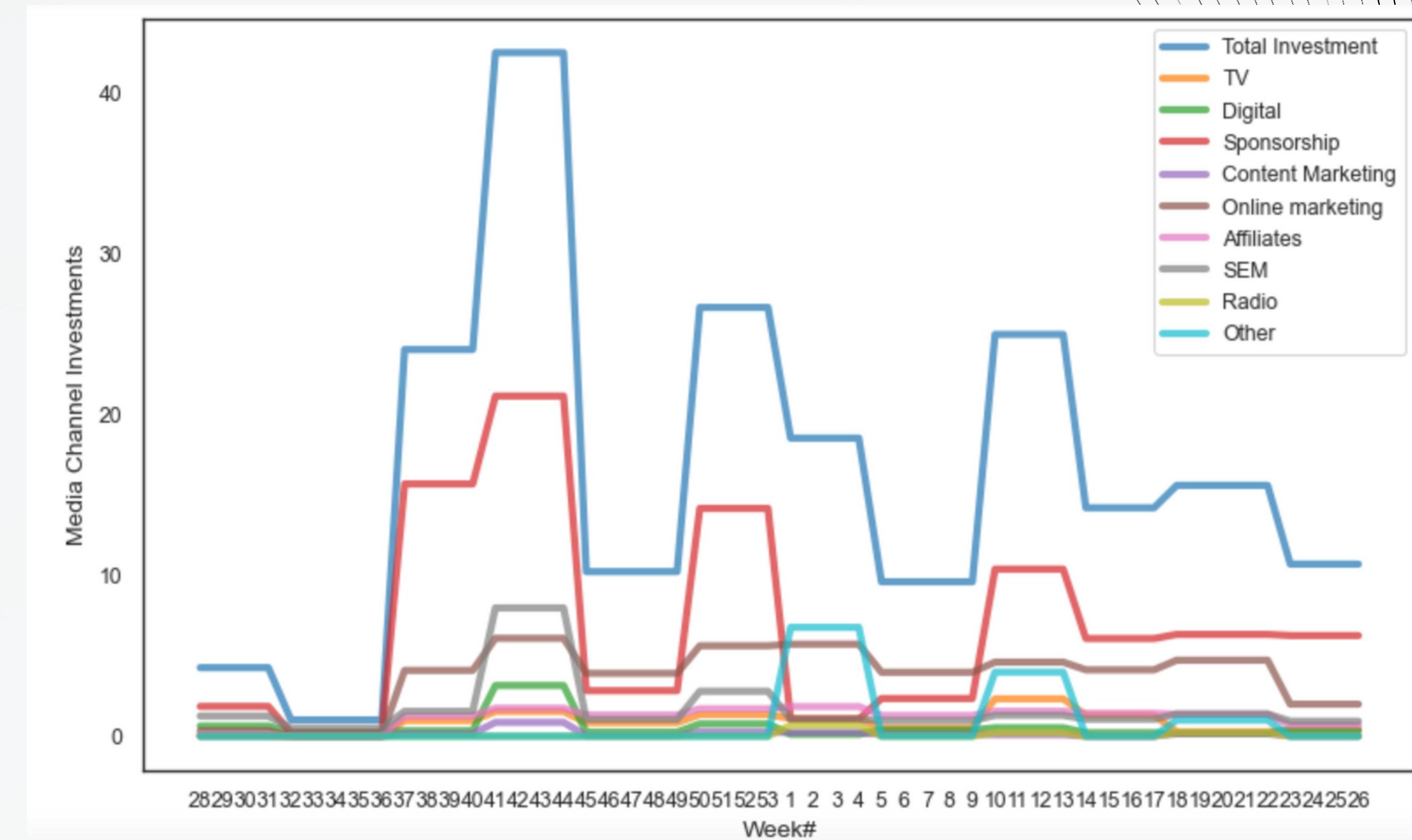
DISTRIBUTION OF DISCOUNT

- The median discount percentage offered for luxury items is less compared to that of Mass Market Products.
- This is a known trend among luxury products or luxury brands to offer limited or no discounts to retain the exclusivity of their products



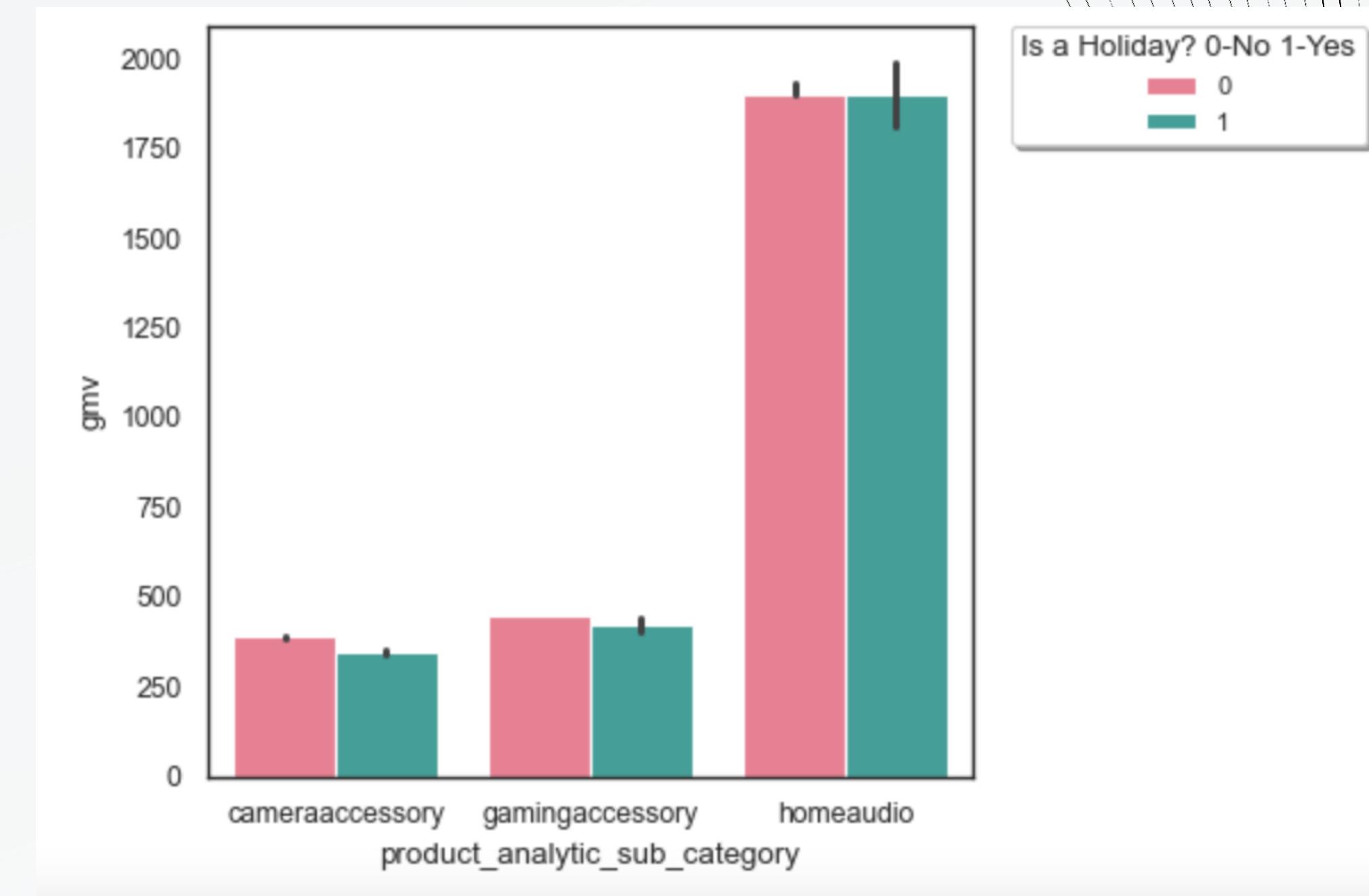
INVESTMENTS ON MEDIA CHANNELS

- Bulk of the Ad Investment has been made in Sponsorships followed by Online Marketing & Search Engine Marketing(specially during Thanksgiving)



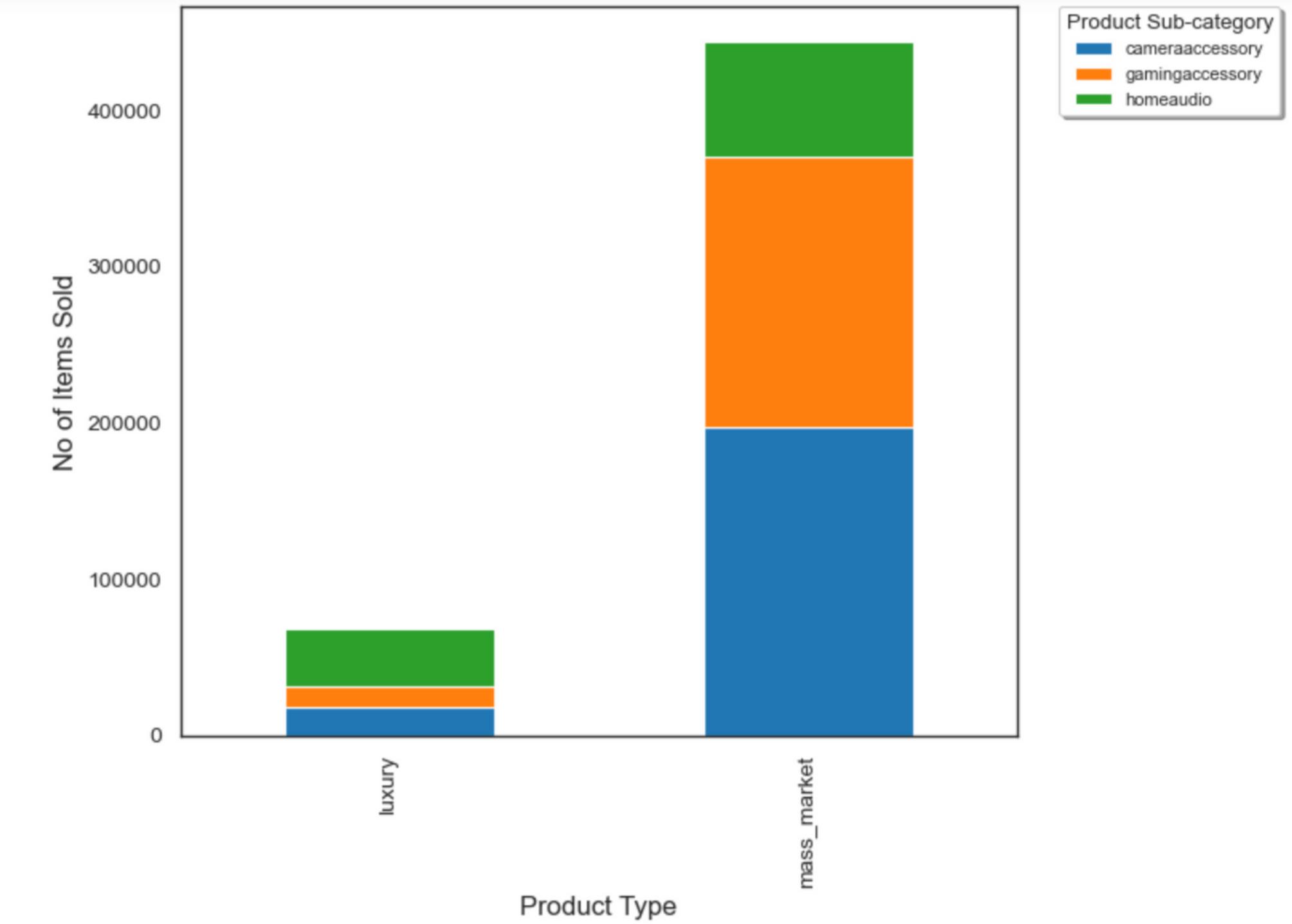
AVERAGE REVENUE FROM HOLIDAYS VS NON - HOLIDAY DAYS

- The average(median) revenue for 3 product sub categories from holiday and non-holiday days are more or less comparable.



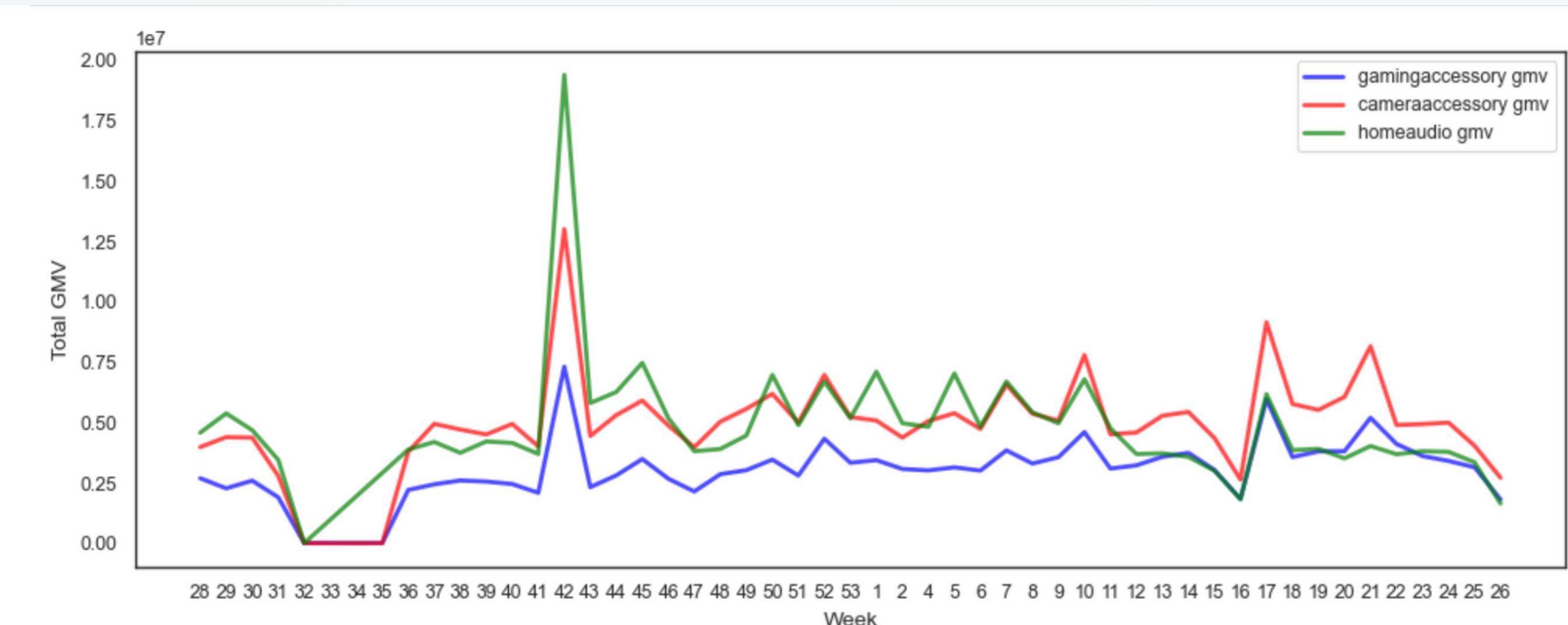
NO OF ITEMS(LUXURY/MASS-MARKET) SOLD

- Most of the units sold belonged to the mass market category.
- Among mass-market products sold, Camera and Gaming Accessories related products were sold the most.
- Home Audio products were the most popular among the luxury products sold



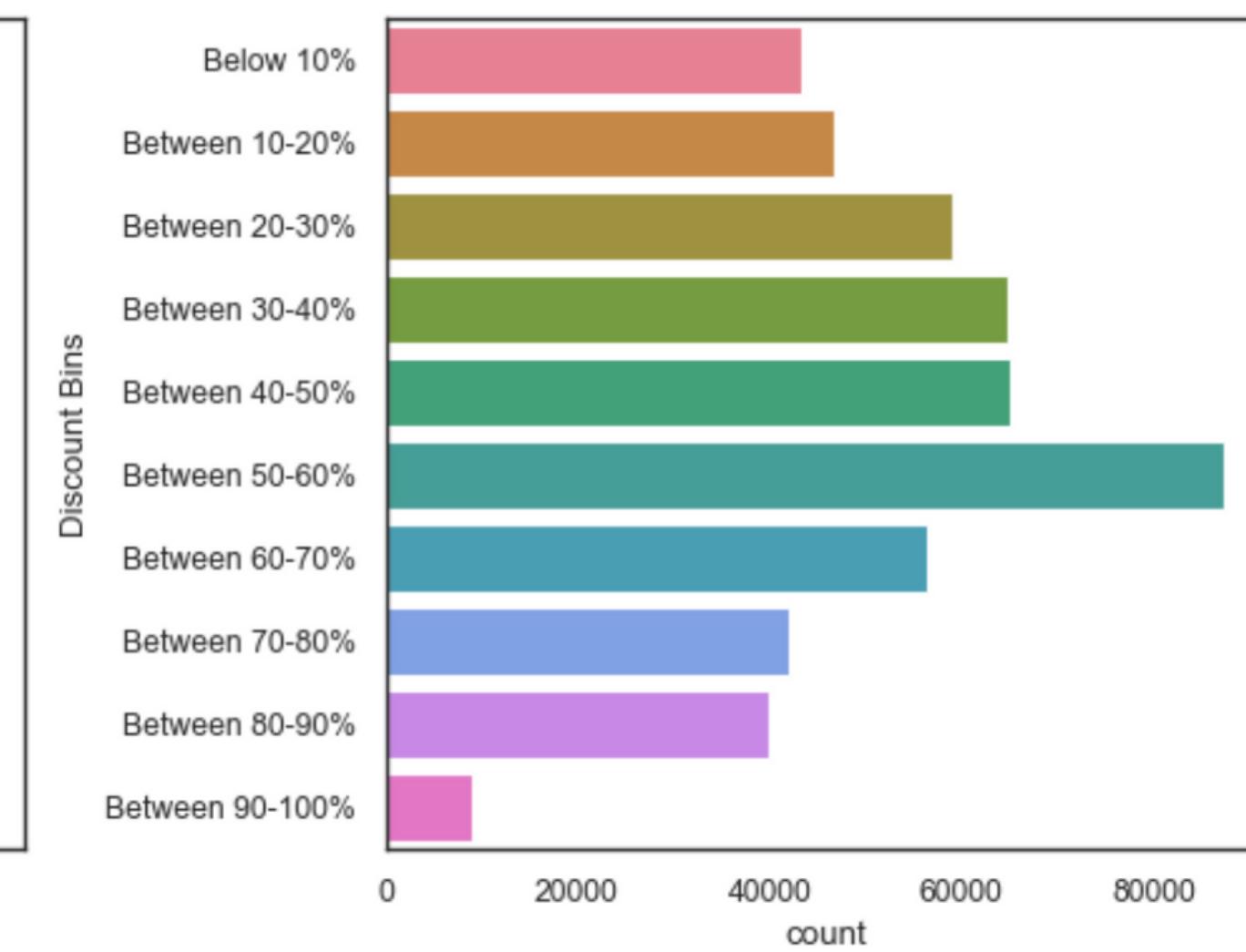
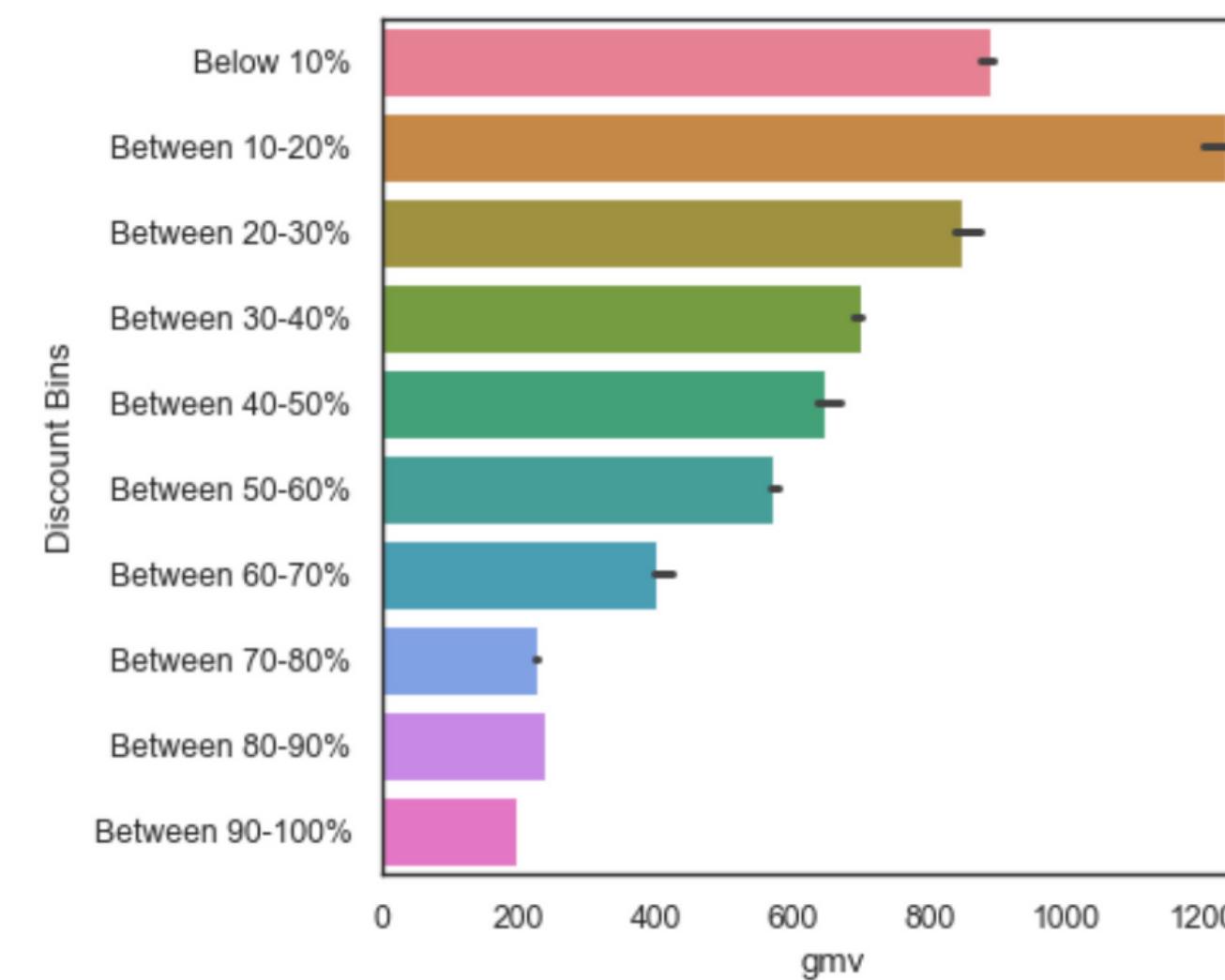
TRENDS OF REVENUE DISCOUNT%

- For the week# 42 (during Thanksgiving), all the graphs show a steep rise. Revenue increased because of both higher discount% and increased Ad Investment.
- In general, the average discount percentage offered for home audio products is lesser compared to that of the other product subcategories.



REVENUE VARIATION WITH DISCOUNT PERCENTAGE

- This shows that at a higher discount, although the sales are good, the revenue collapses signifying a loss for the company. An average discount of 10-20% is the most profitable for the company



LINEAR REGRESSION MODEL FOR CAMERA ACCESSORY(ADDITIONAL MODEL)

- Here the model has a good r2 score which means the model is able to explain more variability.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

cameraaccessory_model = LinearRegression().fit(X_cameraaccessory_train, y_cameraaccessory_train)
y_cameraaccessory_test_pred = cameraaccessory_model.predict(X_cameraaccessory_test)

print('R2 Score: {}'.format(r2_score(y_cameraaccessory_test, y_cameraaccessory_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_cameraaccessory_test, y_cameraaccessory_test_pred)))

R2 Score: 0.822839583484082
Mean Squared Error: 0.17533502288422556
```

LINEAR REGRESSION MODEL FOR GAMING ACCESSORY (ADDITIVE MODEL)

- Here the model has a high r2 score which means the model is able to explain variability much better.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

gamingaccessory_model = LinearRegression().fit(X_gamingaccessory_train, y_gamingaccessory_train)
y_gamingaccessory_test_pred = gamingaccessory_model.predict(X_gamingaccessory_test)

print('R2 Score: {}'.format(r2_score(y_gamingaccessory_test, y_gamingaccessory_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_gamingaccessory_test, y_gamingaccessory_test_pred)))
```

R2 Score: 0.9299819870219652
Mean Squared Error: 0.05603814344437159

LINEAR REGRESSION MODEL FOR HOME AUDIO (ADDITIONAL MODEL)

- Here the model has a really high r2 score which means the model is able to explain variability better than the other two categories.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

homeaudio_model = LinearRegression().fit(X_homeaudio_train, y_homeaudio_train)
y_homeaudio_test_pred = homeaudio_model.predict(X_homeaudio_test)

print('R2 Score: {}'.format(r2_score(y_homeaudio_test, y_homeaudio_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_homeaudio_test, y_homeaudio_test_pred)))
```

R2 Score: 0.9547790089114497

Mean Squared Error: 0.10842933389740747

LINEAR REGRESSION MODEL FOR CAMERA ACCESSORY (MULTIPLICATIVE MODEL)

- Here the model has a good r2 score which means the model is able to explain more variability.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

cameraaccessory_mul_model = LinearRegression().fit(X_cameraaccessory_mul_train, y_cameraaccessory_mul_train)
y_cameraaccessory_mul_test_pred = cameraaccessory_mul_model.predict(X_cameraaccessory_mul_test)

print('R2 Score: {}'.format(r2_score(y_cameraaccessory_mul_test, y_cameraaccessory_mul_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_cameraaccessory_mul_test, y_cameraaccessory_mul_test_pred)))
```

```
R2 Score: 0.8399046760804205
Mean Squared Error: 0.3550539909715613
```

LINEAR REGRESSION MODEL FOR GAMING ACCESSORY (MULTIPLICATIVE MODEL)

- Here the model has a high r2 score which means the model is able to explain variability much better.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

gamingaccessory_mul_model = LinearRegression().fit(X_gamingaccessory_mul_train, y_gamingaccessory_mul_train)
y_gamingaccessory_mul_test_pred = gamingaccessory_mul_model.predict(X_gamingaccessory_mul_test)

print('R2 Score: {}'.format(r2_score(y_gamingaccessory_mul_test, y_gamingaccessory_mul_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_gamingaccessory_mul_test, y_gamingaccessory_mul_test_pred)))
```

R2 Score: 0.9064990493438985
Mean Squared Error: 0.1469344986154294

LINEAR REGRESSION MODEL FOR HOME AUDIO (MULTIPLICATIVE MODEL)

- Here R2 is negative which signifies that the chosen model does not follow the trend of the data, so fits worse than a horizontal line. It simply means the chosen model (with its constraints) fits the data really poorly.

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error

homeaudio_mul_model = LinearRegression().fit(X_homeaudio_mul_train, y_homeaudio_mul_train)
y_homeaudio_mul_test_pred = homeaudio_mul_model.predict(X_homeaudio_mul_test)

print('R2 Score: {}'.format(r2_score(y_homeaudio_mul_test, y_homeaudio_mul_test_pred)))
print('Mean Squared Error: {}'.format(mean_squared_error(y_homeaudio_mul_test, y_homeaudio_mul_test_pred)))
```

R2 Score: -0.6487144038500239

Mean Squared Error: 0.34500556402491367

MODEL SELECTION

MODEL SELECTION FROM ADDITIVE, MULTIPLICATIVE, KOYCK, DISTRIBUTIVE LAG MODEL (ADDITIVE), DISTRIBUTIVE LAG MODEL (MULTIPLICATIVE)

- We notice that all the 3 chosen models for the 3 sub-categories are Multiplicative models. This fact tells us that there exists some interaction between the KPIs. These models tell us about the growth of revenue vs the interactive growth of the KPIs.

Product Sub-category	Linear Regression Model	R-square on Test Dataset	Mean Square Error	Top 5 KPIs
cameraaccessory	Multiplicative with CV	0.91	0.09	product_vertical_lens (0.181) product_vertical_camerabattery (0.160) is_mass_market (0.149)
gamingaccessory	Multiplicative with CV	0.94	0.06	product_vertical_camerabatterycharger (0.121) TV (0.105) product_vertical_gamingheadset (0.250) is_mass_market (0.234) product_vertical_gamingmouse (0.224)
cameraaccessory	Multiplicative with CV	0.86	0.14	product_vertical_homeaudiospeaker (0.469) is_mass_market (0.289) product_vertical_fmradio (0.224) Radio_Ad_Stock (0.147) Sponsorship (0.121)

RECOMMENDATIONS

- The company should promote Lens, Camera Batteries, and Camera Battery Chargers as they fetch the highest revenue.
- The company should promote Gaming Headset, Gaming Mouse, and Gamepad as they fetch the highest revenue. On the contrary, Gaming Memory Cards result in loss.
- The company should promote Home Audio Speakers and FM Radios as they fetch the highest revenue.
- EDA shows that an average discount percentage between 10-20% is the most profitable for the company, especially among luxury items.



THANK YOU!