MARKETING AND RETAIL ANALYTICS

ANALYTICS

PROJECT

The Baking Hub

BUSINESS PROBLEM

"Pancake Mixes" and "Syrups" category contributes only 5.6% and 19.1% of total Sales respectively hence, identification of opportunities to drive growth by increasing household penetration into 'syrups' and 'pancake mixes' category.

STEP 1: Segment Migration Analysis

Row Label	Monetray	Frequency	Recency	Variety	Recency Score	R score	F score	M Score	V Score	Total Score	
1	6.47	4	345	2	19	5	5	4	5	19	Segment 4
2	38.99	11	355	2	9	5	5	5	5	20	Segment 4
3	5.28	2	322	2	42	4	4	4	5	17	Segment 4
5	3.75	2	71	2	293	1	4	3	5	13	Segment 3
7	1.69	1	25	1	339	1	3	1	4	9	Segment 2
11	3	1	204	1	160	3	3	3	4	13	Segment 3
12	2.09	1	323	1	41	4	3	2	4	13	Segment 3
13	2.99	1	315	1	49	4	3	2	4	13	Segment 3
22	5.97	2	152	1	212	2	4	4	4	14	Segment 3
23	1.79	1	97	1	267	1	3	1	4	9	Segment 2
24	2.29	1	336	1	28	5	3	2	4	14	Segment 3
25	2.29	1	7	1	357	1	3	2	4	10	Segment 2
26	4.98	2	357	2	7	5	4	4	5	18	Segment 4
29	4.78	2	23	2	341	1	4	4	5	14	Segment 3
31	3.49	1	74	1	290	1	3	3	4	11	Segment 2
32	0.99	1	286	1	78	4	3	1	4	12	Segment 2
34	2.39	1	307	1	57	4	3	2	4	13	Segment 3
36	17.94	6	301	1	63	4	5	5	4	18	Segment 4
39	5.27	3	143	1	221	2	5	4	4	15	Segment 3
41	2.25	1	29	1	335	1	3	2	4	10	Segment 2
44	1.19	1	263	1	101	4	3	1	4	12	Segment 2
46	2.39	1	350	1	14	5	3	2	4	14	Segment 3

	R ranks	F ranks	M ranks	V ranks
20%	41	1	1.99	
40%	102	1	3	
60%	174	2	4.68	
80%	245	3	8.17	
100%	363	111	490.81	

STEP 1 OUTPUT R-F-M-V score for Year-1.

Dataset Used- transaction.csv

Tools Used
- Google Collab
- Excel
- SPSS

We then assigned each customer to a segment based on their RFMV score, with Segment 4 representing the high-value customers and Segment 1 representing the low-value customers.

Row Labels	Monetray	Frequency	Recency	Variety	Recency Score	R score	F score	M Score	V Score	Total Score	
1	5.83	4	700	2	28	5	5	4	5	19	Segment 4
2	50.63	17	699	2	29	5	5	5	5	20	Segment 4
8	1.99	1	377	1	351	1	3	2	4	10	Segment 2
10	7.37	3	428	2	300	1	5	4	5	15	Segment 3
14	3.49	1	379	1	349	1	3	3	4	11	Segment 2
39	3.98	2	505	1	223	2	4	3	4	13	Segment 3
49	14.3	5	651	1	77	4	5	5	4	18	Segment 4
50	1.59	1	521	1	207	2	3	1	4	10	Segment 2
51	2.62	2	448	2	280	1	4	2	5	12	Segment 2
52	3.66	2	543	1	185	2	4	3	4	13	Segment 3
54	1.5	1	529	1	199	2	3	1	4	10	Segment 2
55	1.37	1	574	1	154	3	3	1	4	11	Segment 2
57	1.99	1	531	1	197	2	3	2	4	11	Segment 2
62	2.99	1	378	1	350	1	3	3	4	11	Segment 2

	R ranks	F ranks	M ranks	V ranks	
20%	40	1	1.99	1	
40%	100	1	2.99	1	
60%	169	2	4.64	1	
80%	241	3	8.04	2	
100%	363	115	614.81	2	

Our exploratory and descriptive analysis identified several key findings: Segment migration analysis showed that 32% of customers churned in the second year, with high-value segment customers showing higher churn rates.

Migration ananlysis from Year 1 to Year 2 based on RFMV segmentation

Count of Customer ID	Column Labels	Customer Segmentation Year 2			
Row Labels	Lost Customer	Segment 2	Segment 3	Segment 4	Grand Total
New Customer		43567	38744	18251	100562
Segment 2	43778	9305	8485	5785	67353
Segment 3	33239	9738	10718	9353	63048
Segment 4	13497	6622	10278	22037	52434
Grand Total	90,514.00	69,232.00	68,225.00	55,426.00	2,83,397.00

Customer Segmentation Year 1

Count of Customer ID	Column Labels	Customer Segmentation Year 2			
Row Labels	Lost Customer	Segment 2	Segment 3	Segment 4	Grand Total
New Customer	0.00%	15.37%	13.67%	6.44%	35.48%
Segment 2	15.45%	3.28%	2.99%	2.04%	23.77%
Segment 3	11.73%	3.44%	3.78%	3.30%	22.25%
Segment 4	4.76%	2.34%	3.63%	7.78%	18.50%
Grand Total	31.94%	24.43%	24.07%	19.56%	100.00%

Customer Segmentation Year 1

Total Positive change	43.82%
Total Negative change	41.34%
Total Netural	14.84%
Total %	100.00%

STEP 2: Price Elasticity Analysis

Dataset Used-

- **transaction.csv**
- **causal.csv**
- **product.csv**

Categorized the data by different commodities such as pasta sauce, pasta, pancake mix, and syrup.To calculate the elasticity for each commodity, we first obtained the product sales data and price information from the store's transaction records.

We then used statistical models to analyze the relationship between price and sales volume for each commodity. We calculated the percentage change in sales volume for each 1% change in price, which gave us a measure of the product's price elasticity.

```python
# Define the independent and dependent variables
X1 = df.drop(['Row Labels', 'ln(sales)'], axis=1)
y1 = df['ln(sales)'] # dependent variable

# Add constant to the independent variables
X1 = sm.add_constant(X1)

# Fit the linear regression model
model1 = sm.OLS(y1, X1).fit()

# Print the model summary
print(model1.summary())
```

	coef	std err	t	P
const	8.6673	0.190	45.723	0
ln(baseprice)	-0.3934	0.338	-1.165	0
ln(api)	-0.1676	0.410	-0.409	0
sum of feature_desc_Back Page Feature	0.0001	0.001	0.122	0
Sum of feature_desc_Front Page Feature	1.155e-13	1.6e-14	7.200	0
Sum of feature_desc_Interior Page Feature	0.0011	0.023	0.046	0
Sum of feature_desc_Interior Page Line Item	4.675e-14	4.09e-15	11.427	0
Sum of feature_desc_Not on Feature	-0.0002	0.005	-0.045	0
Sum of feature_desc_Wrap Back Feature	-4.134e-14	3.54e-15	-11.661	0
Sum of feature_desc_Wrap Front Feature	-4.32e-05	0.004	-0.011	0
Sum of feature_desc_Wrap Interior Feature	3.716e-15	9.63e-16	3.861	0

For determining price-elasticity for pancakes, similar process was followed for syrup.

Impact on sales

Also the promotional price elasticity is
very poor even after promotion the
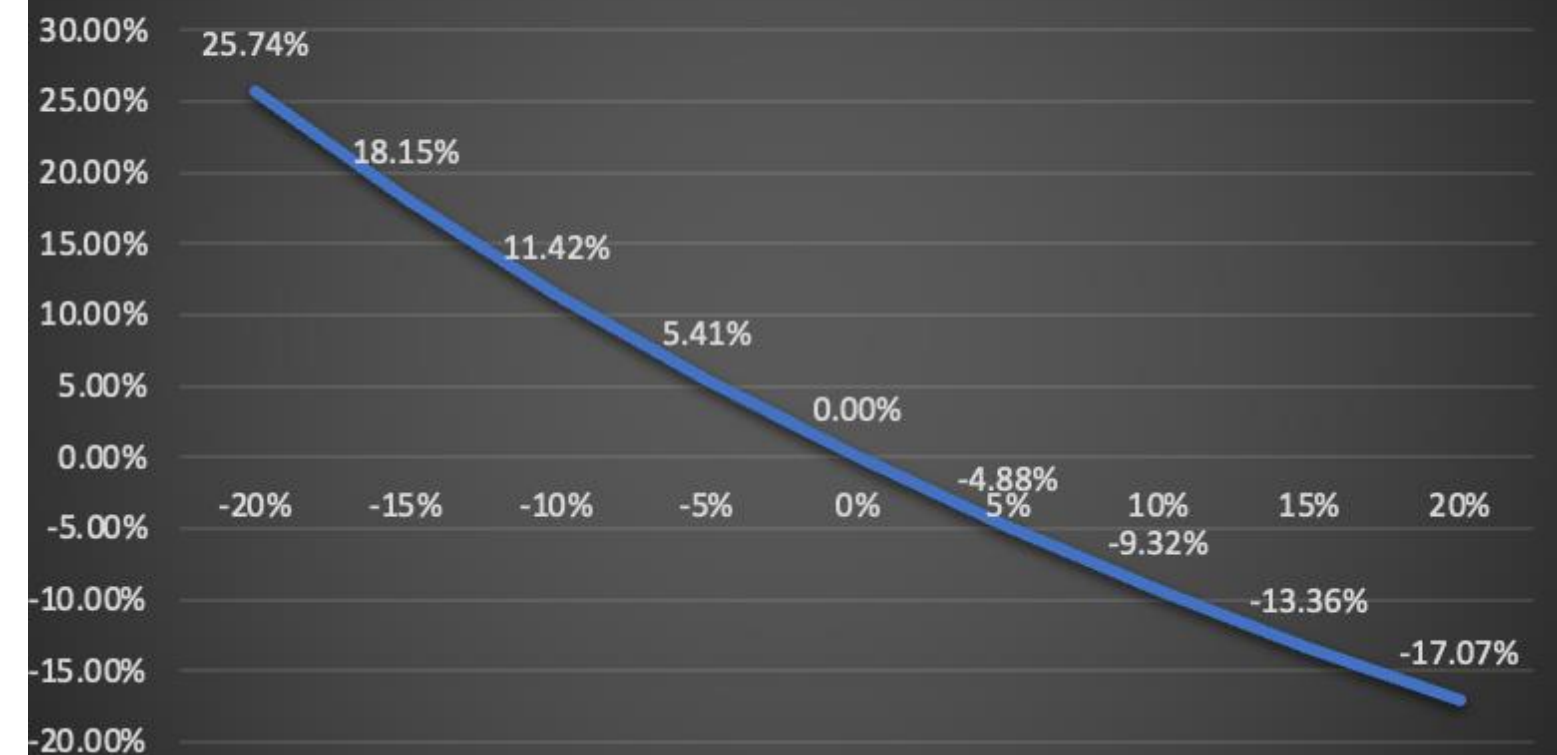product is not getting sold
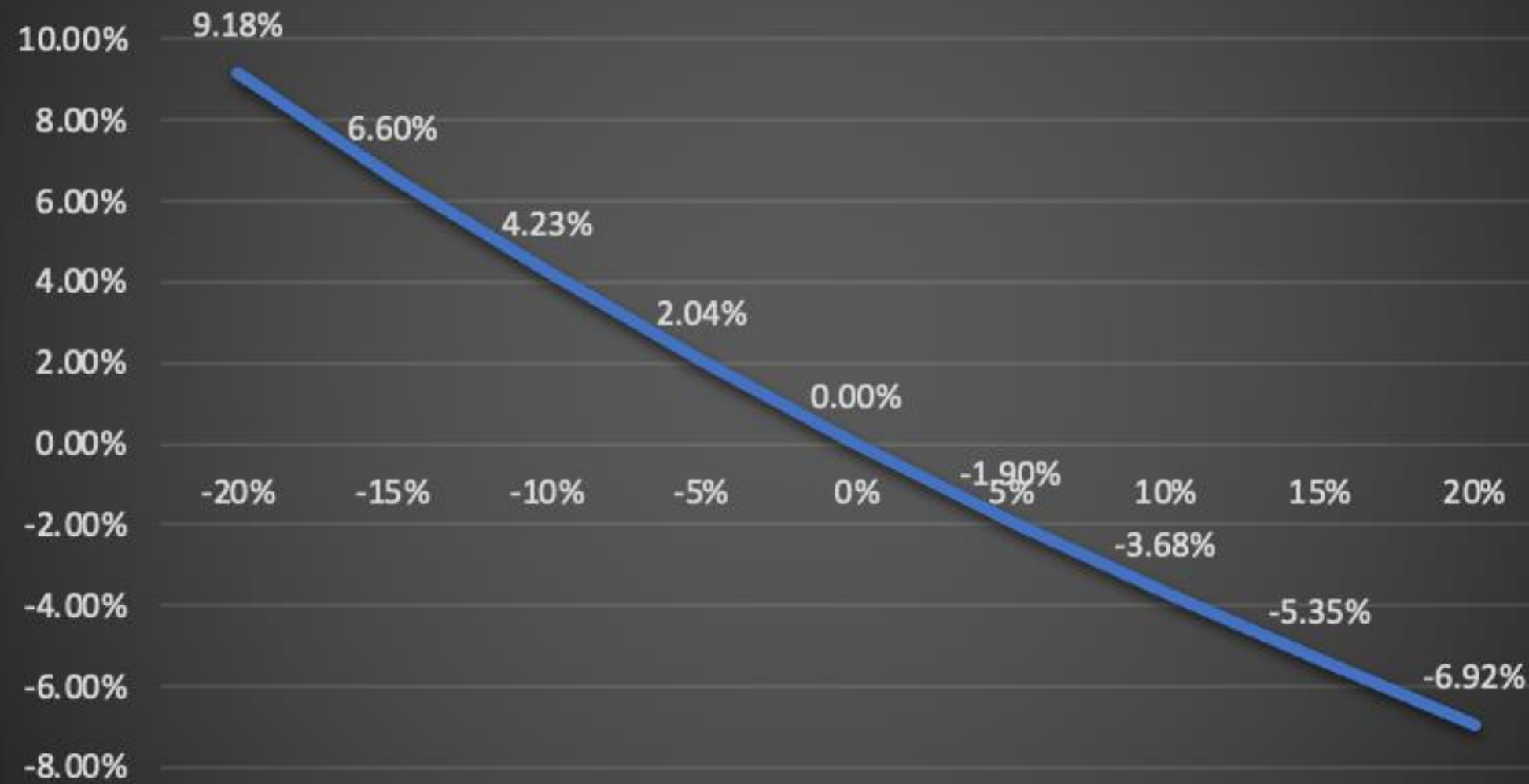
BASE PRICE ELASTICITY

Compared to other commodities pancakes have highest base price elasticity, means if price changes there is significant effect on sales

Base Price Elasticity Curve Table	
Base Price elasticity	-0.3934
% change in price	Impact on sales
-20%	9.18%
-15%	6.60%
-10%	4.23%
-5%	2.04%
0%	0.00%
5%	-1.90%
10%	-3.68%
15%	-5.35%
20%	-6.92%

Impact on sales

PROMOTIONAL PRICE ELASTICITY

Impact on sales
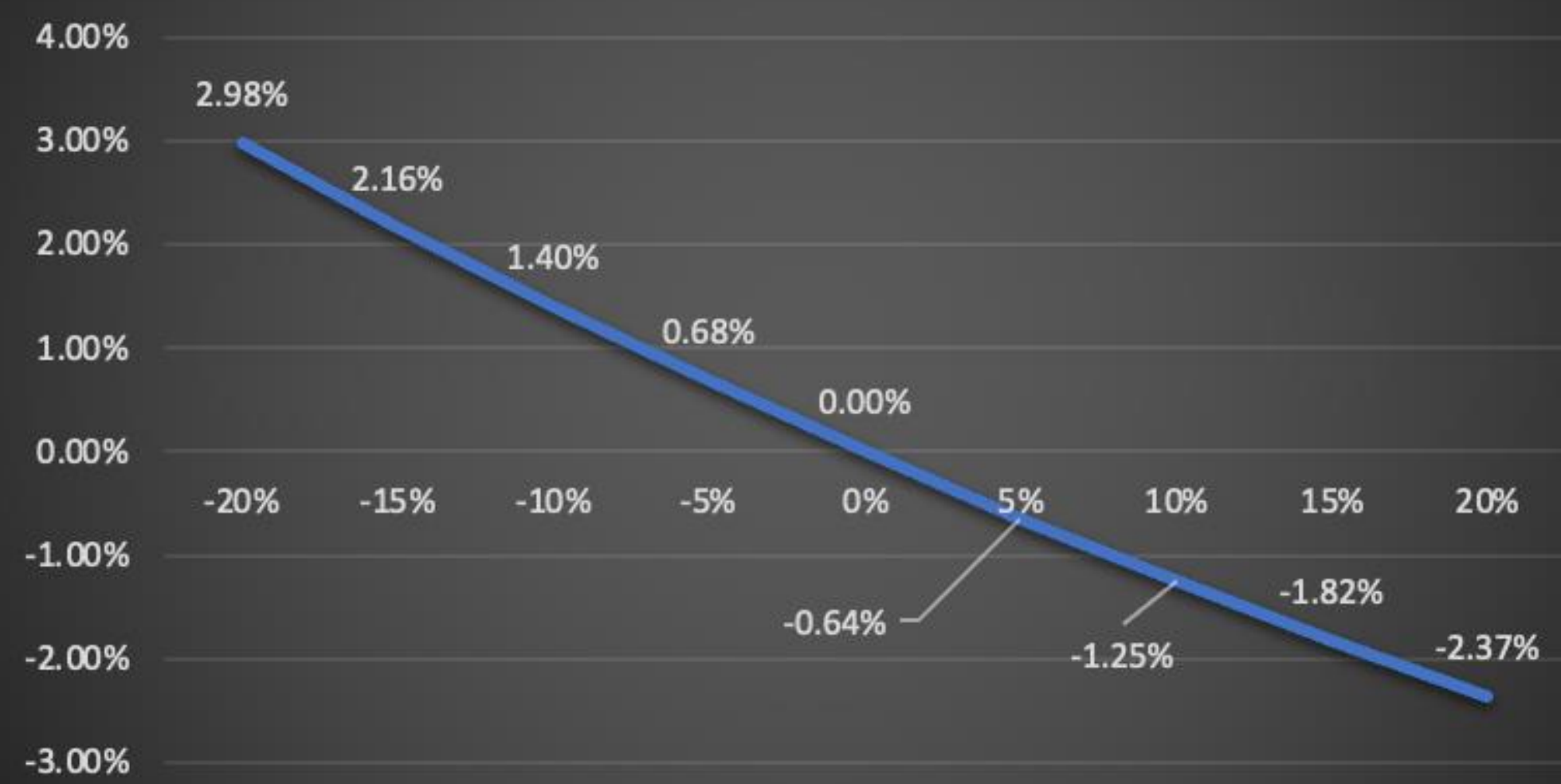
There is significant impact on sales in case of syrup, so it might be promotion is not at optimum level which might be hindrance in sales
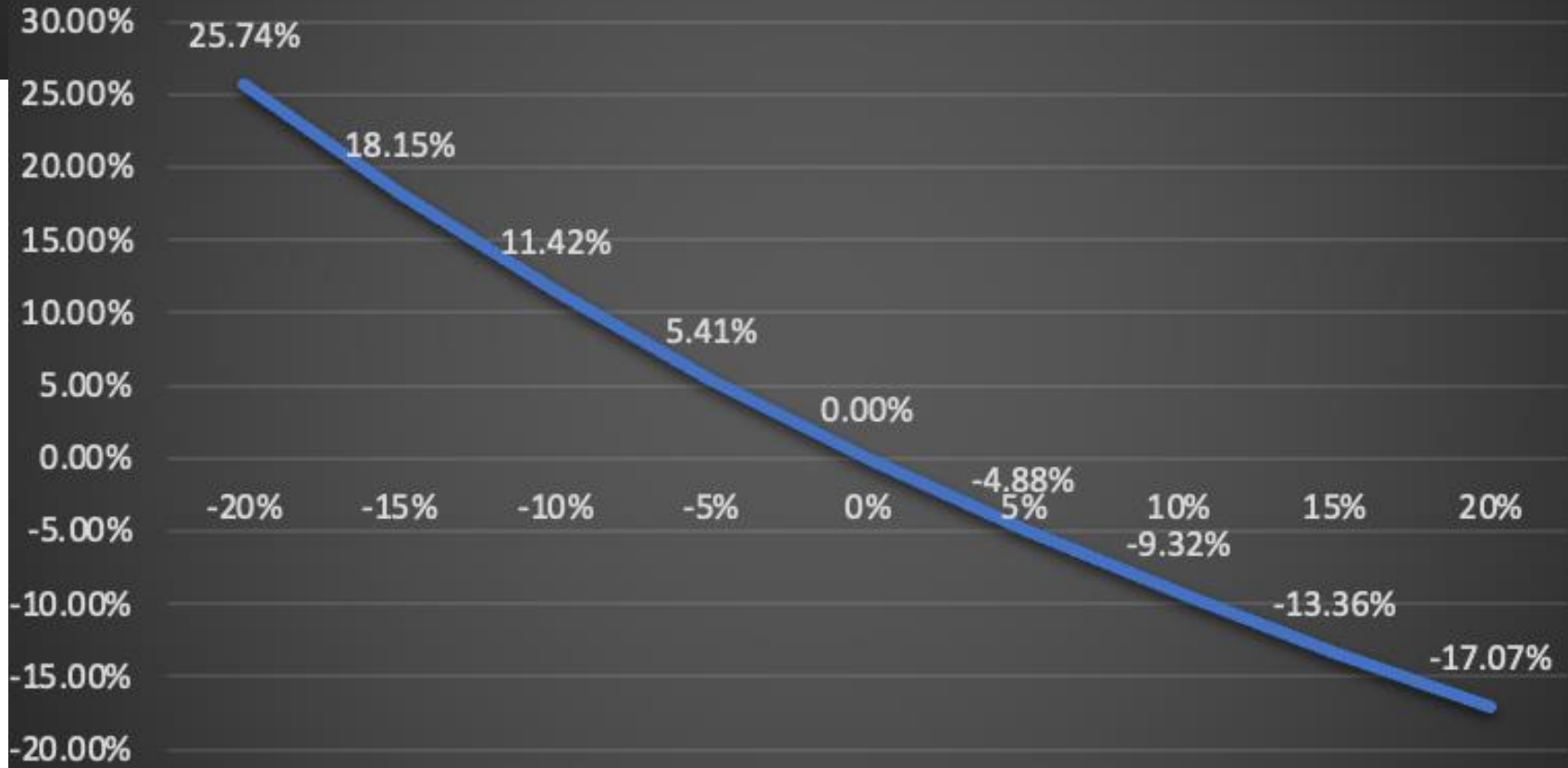
BASE PRICE ELASTICITY

Syrup has low elasticity but not as low as pasta sauce

Base Price Elasticity Curve Table	
Base Price elasticity	-0.1316
% change in price	Impact on sales
-20%	2.98%
-15%	2.16%
-10%	1.40%
-5%	0.68%
0%	0.00%
5%	-0.64%
10%	-1.25%
15%	-1.82%
20%	-2.37%

Impact on sales

PROMOTIONAL PRICE ELASTICITY

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	−5481.818	5938.106		−0.923	0.358
	Base	1.188	0.132	0.653	9.032	0.000
	Baseline price	2039.305	2518.585	0.064	0.810	0.420
	Feature	5.995	44.131	0.018	0.136	0.892
	Display	24.603	9.716	0.361	2.532	0.013
	Dummy	−591.164	332.326	−0.191	−1.779	0.078
a. Dependent Variable: Sum of units						

Sales Volume Decomposition, First year

Effectiveness of last year

Sales Volume Decomposition, Last 52 Weeks
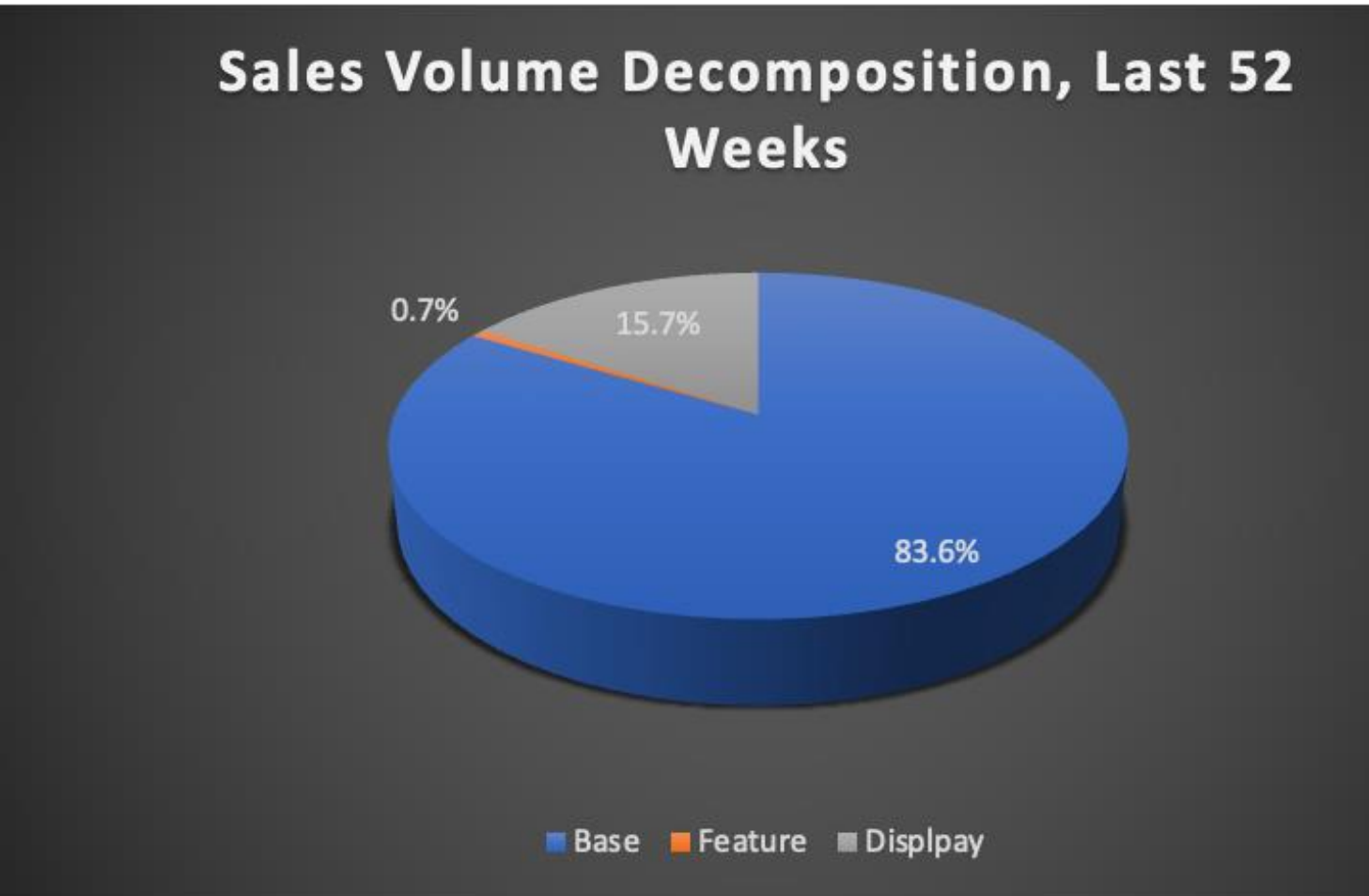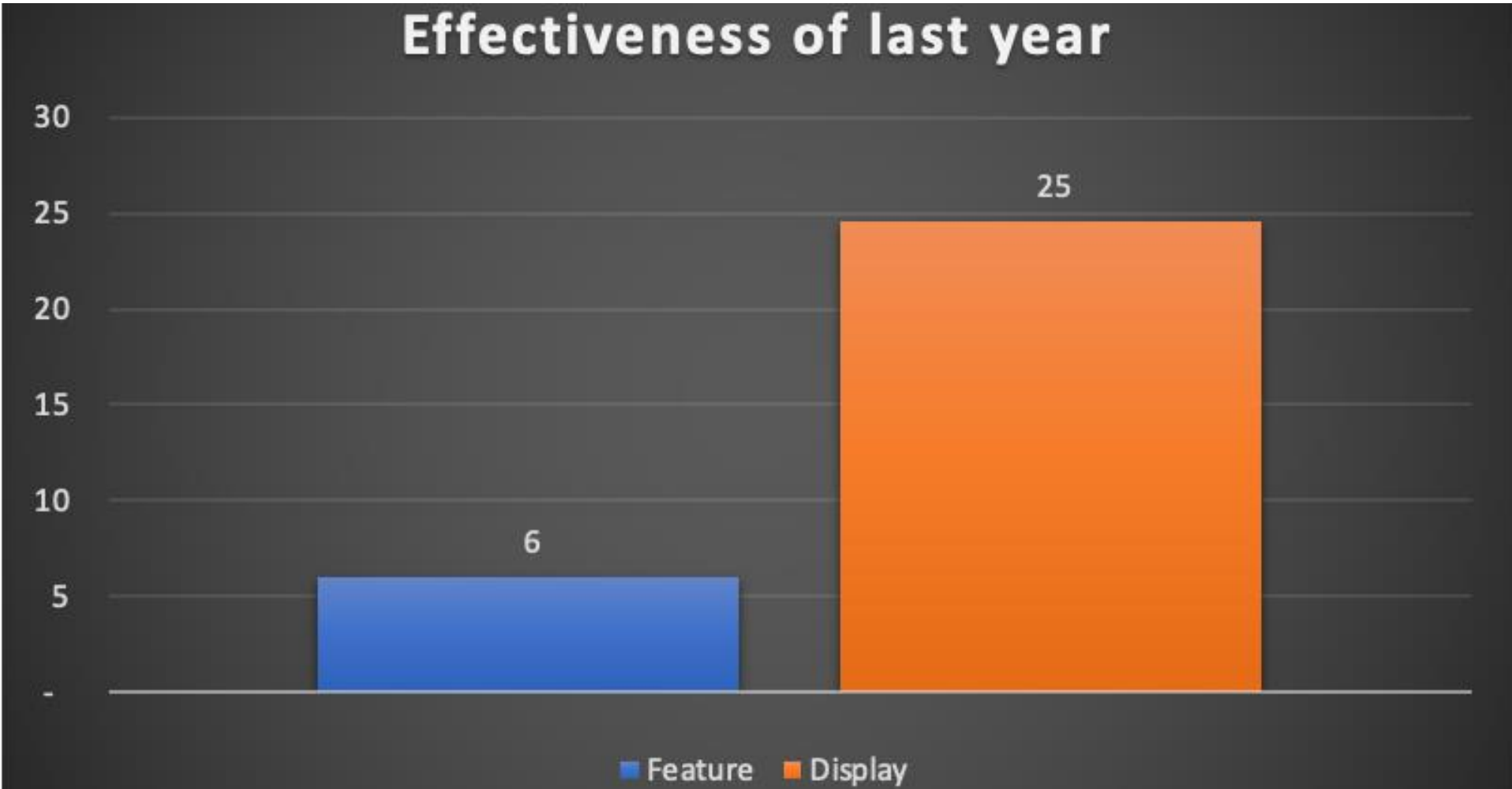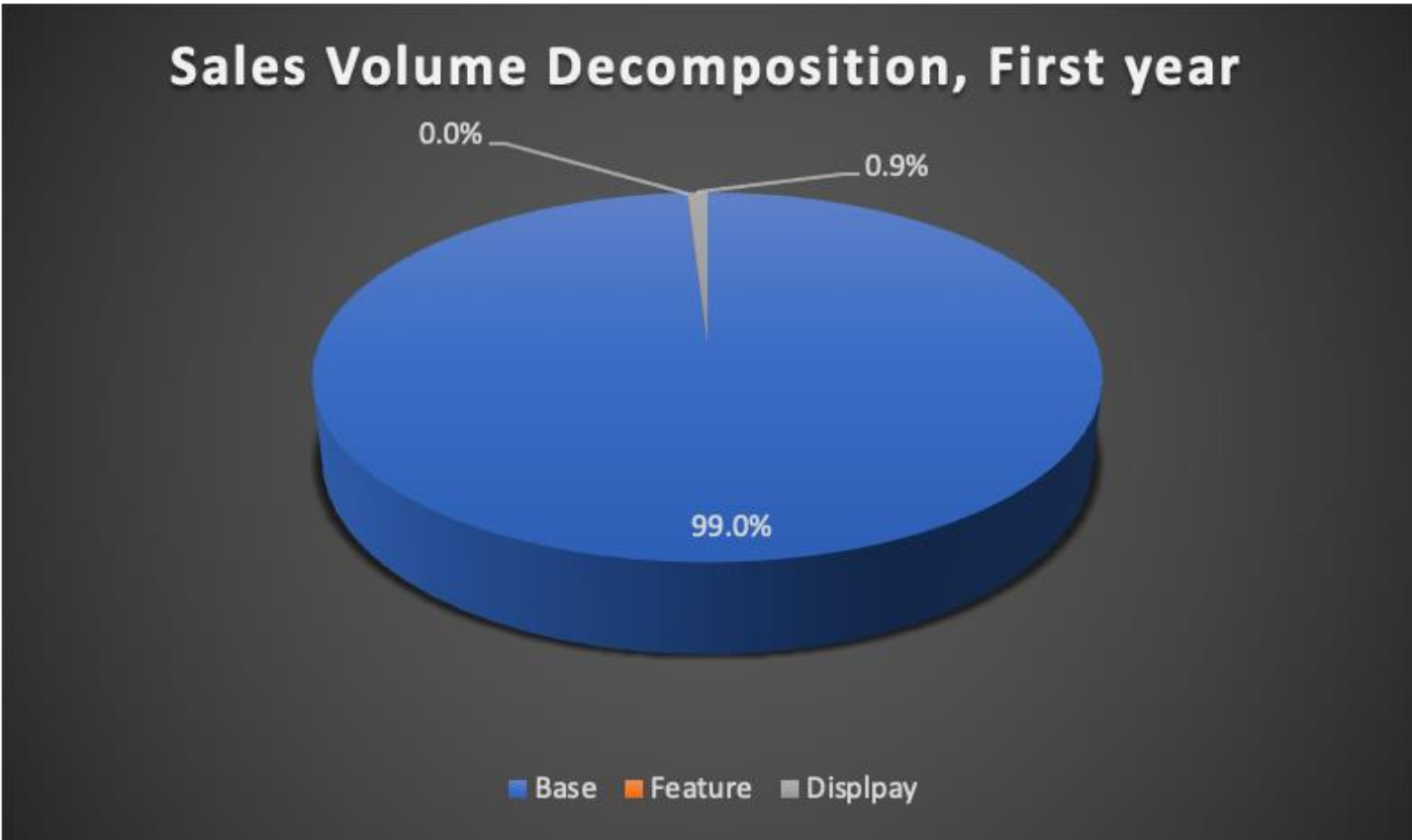
Insights and recommendations

Optimize promotional strategies: The store's **promotional strategies were not effective**, and promotional activity was low in year one. To increase revenue, the store should try different promotional strategies and make them more effective, with a focus on display.

Make **products less price sensitive**: Our price elasticity analysis showed that the store's products were price-sensitive. To reduce customer churn and increase revenue, the store should make products less price-sensitive through horizontal expansion, vertical stretch of segmentation, and promoting premium segments.

Retention strategies: The company **needs to focus on retaining its high value customers**, as they account for the majority of its revenue. This can be achieved by offering better products and services, and providing personalized promotions and discounts.