GROUP – 36

BUAN 6337.501

Submitted By:

Arshiya Anjum Yerur (axy220012)

Bala Krishna Bobbili (bxb220009)

Sneha Thunga (sxt220053)

Sri Dharshan Vasan Senthil Kumaran (sxv220012)

DATA DESCRIPTION AND PROPOSAL

- IRI's Point of Sales: Dataset of IRI Standard Regions and MULO+C
- IRI's Panel: Network panelists data who are representative of US Population
- NPD'S National Eating Trends: Datasets collected from the consumption habits of consumers.
- Analysis is done on all the different categories to understand market trends.
- The project is mainly focused on the analysis and steps to improve the Margarine sales taken by joining the IRI's Table spreads and the Table spread attribute datasets.
- There are 25 variables in the dataset.
- After the data cleaning, the dataset has 37732 records.

MARKET ANALYSIS

From the Market research,

- Imperial Vegetable Oil spread Tub has the least price per OZ.

- Kerrygold Grass-Fed Pure Irish Salted Butter sticks is the expensive product among the various products.

MARKET ANALYSIS

As per customer reviews,

- Overall, the "Country Crock" brand is the dominating brand with the highest sales recorded.

- Highest sales recorded for Original Vegetable Oil with 31723 reviews.

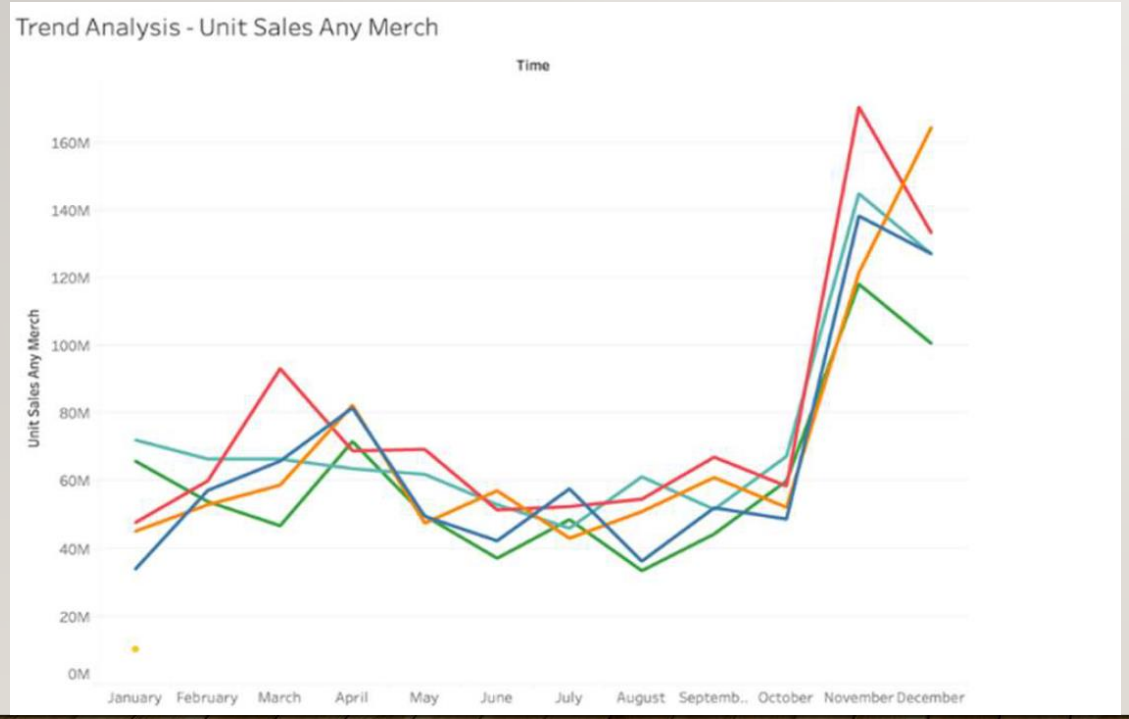
DATA ANALYSIS - DESCRIPTIVE

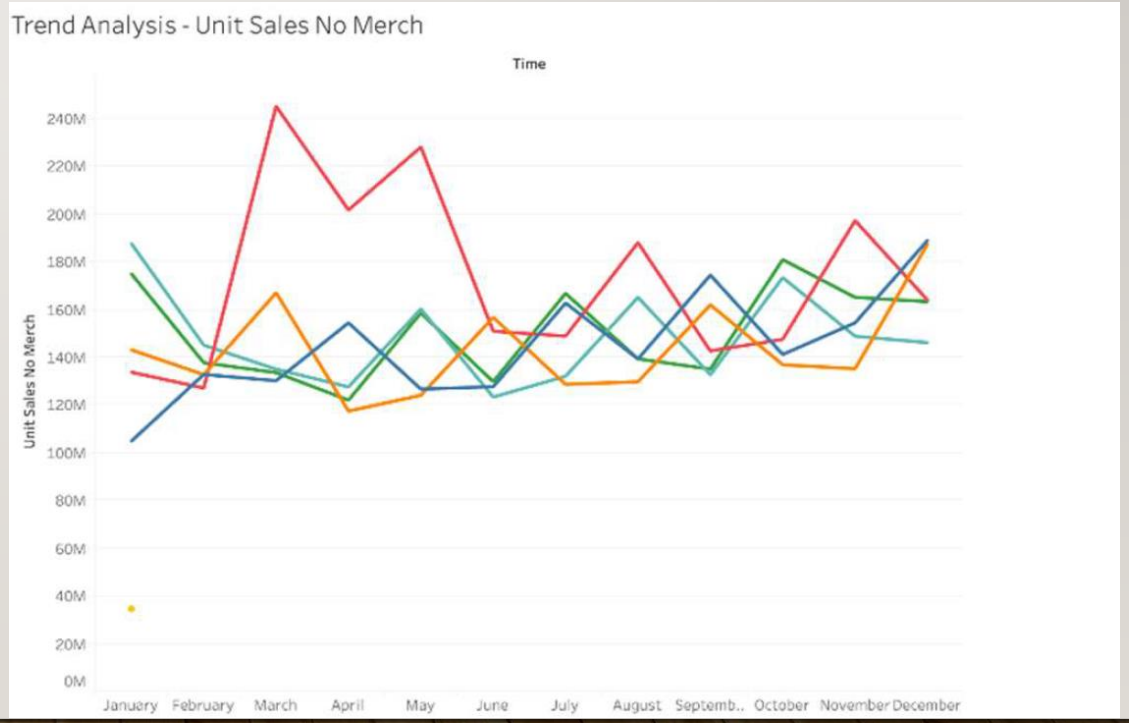
- Summary statistics for sales collected through 2018-2022 for each Geographical location.

	Total Sales							
Geography	Count	Mean	Std. Dev.	Min	25%	50%	75%	Max
California - IRI Standard - Multi Outlet + Conv	56,801	33,386	74,300.84	1.88	2,509	8,691	33,565	2,696,891.28
Great Lakes - IRI Standard - Multi Outlet + Conv	78,452	38,235	107,574.01	1.51	1,404	7,008	31,012	3,511,234.61
Mid-South - IRI Standard - Multi Outlet + Conv	82,852	30,622	83,468.52	2.06	1,430	6,206	23,921	2,717,802.66
Northeast - IRI Standard - Multi Outlet + Conv	85,984	46,743	122,324.94	1.82	2,471	10,818	39,613	4,951,354.24
Plains - IRI Standard - Multi Outlet + Conv	70,679	20,402	53,321.51	2.84	996	4,201	16,192	1,258,688.87
South Central - IRI Standard - Multi Outlet + Conv	59,621	33,894	76,237.56	1.29	1,774	7,470	28,982	1,924,306.38
Southeast - IRI Standard - Multi Outlet + Conv	66,295	44,294	109,593.71	2.12	2,322	11,248	41,596	4,219,204.05
Total US - Multi Outlet + Conv	181,791	115,066	373,285.61	1.08	3,094	13,862	74,522	17,570,600.31
West - IRI Standard - Multi Outlet + Conv	67,761	33,939	74,417.17	2.50	1,544	6,921	29,639	1,561,605.08

DATA ANALYSIS - DESCRIPTIVE

- Trend analysis of Unit sales Any Merch and Unit sales No Merch





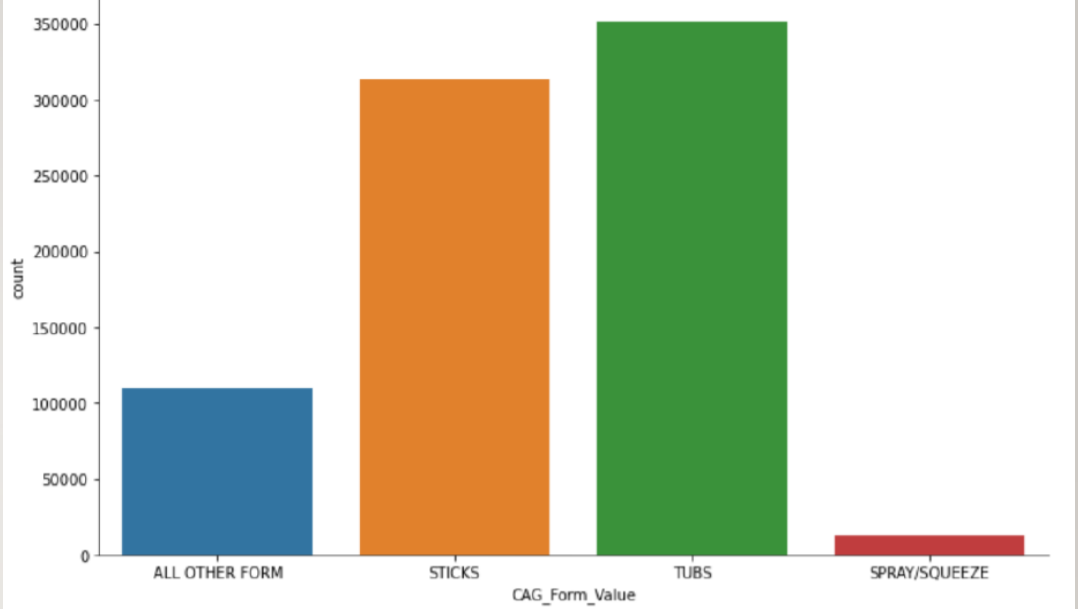
HYPOTHESIS

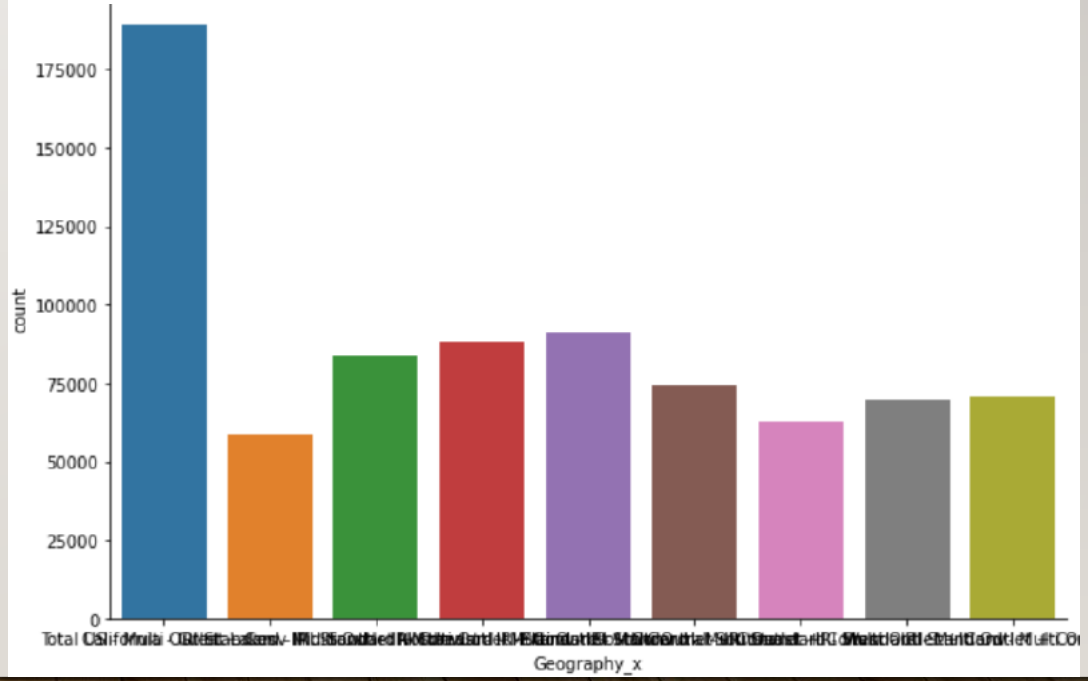
ANOVA TEST FOR FORM AND TOTAL SALES:

- There is a significant difference between the means of total sales of all different forms.

ANOVA TEST FOR GEOGRAPHY AND TOTAL SALES:

- There is a significant difference between the means of total sales in 8 geographical locations.





HYPOTHESIS

CHI SQUARE TEST FOR GEOGRAPHY AND FORM:

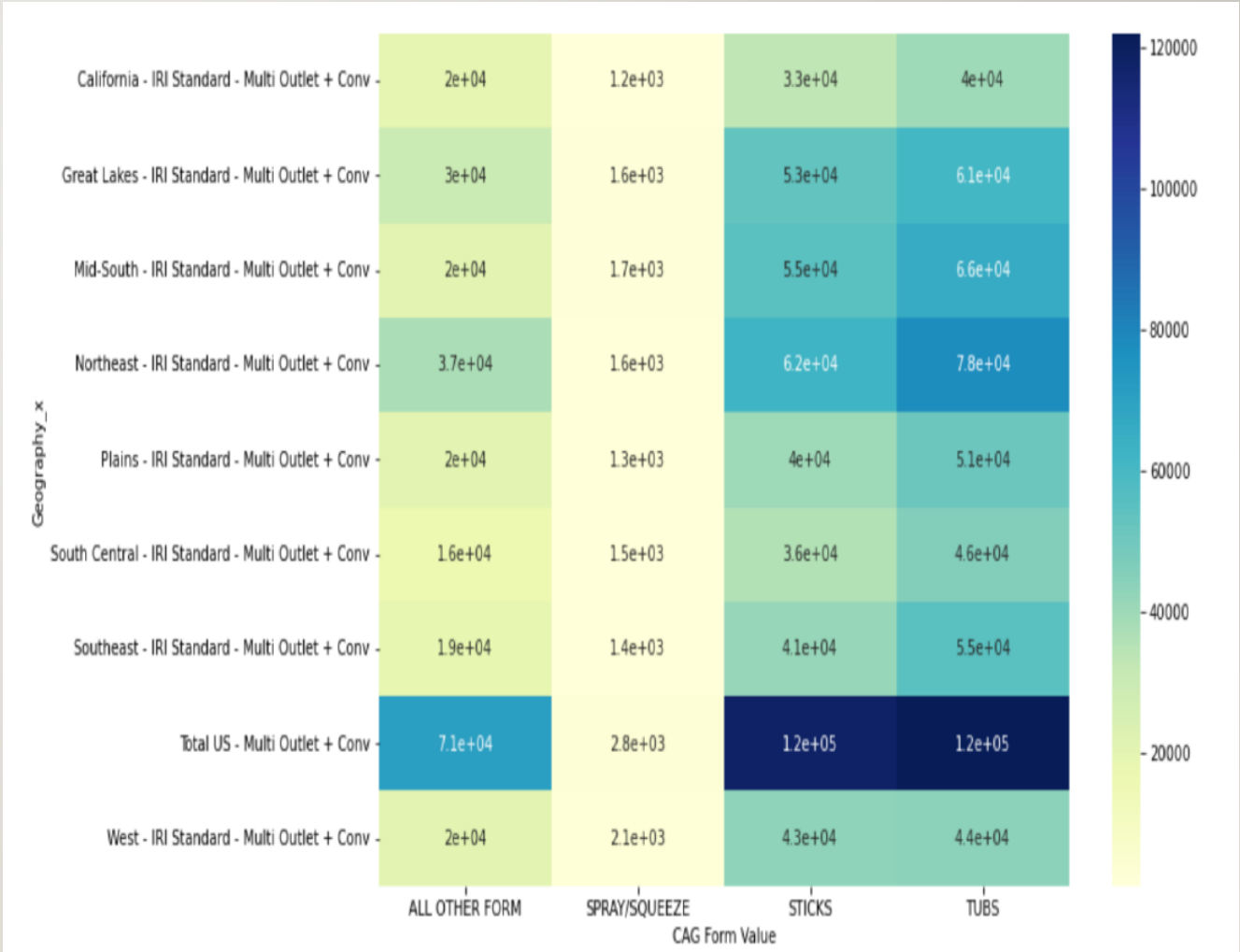
- Geography and form are dependent.

CORRELATION ANALYSIS BETWEEN TABLE SPREADS, COOKING AND SALAD OILS, COOKING SPRAY:

- The Sales of oils/sprays affect the sales of Table spreads.

PRICE OPTIMIZATION TO INCREASE SALES:

- Optimizing the product price of margarine will increase the dollar sales.

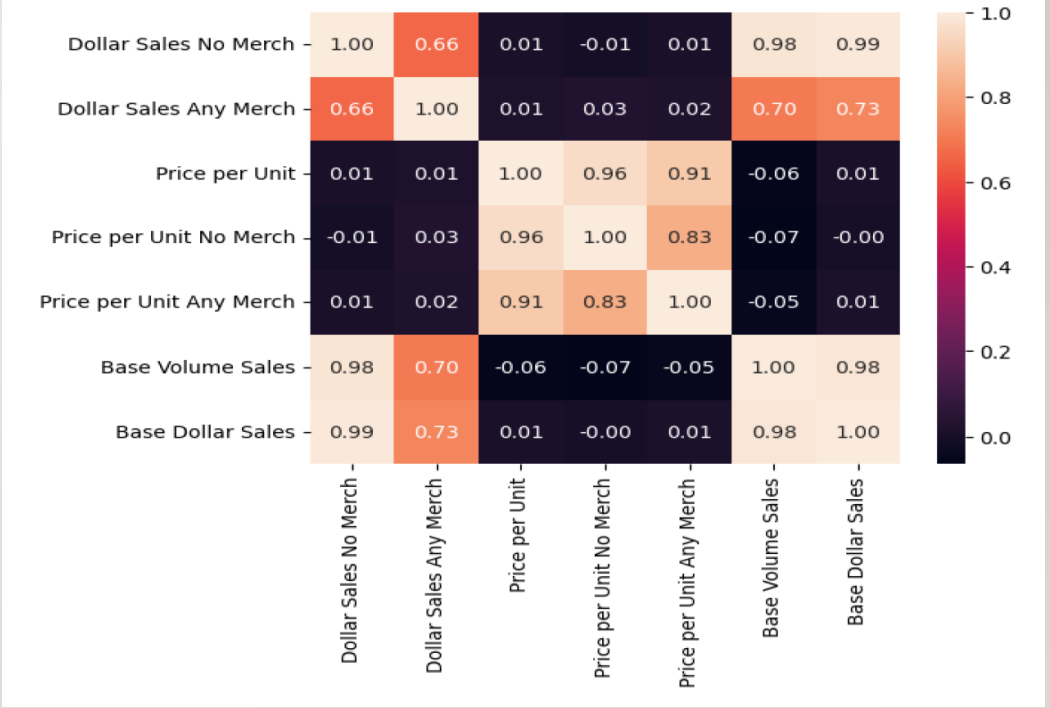


REGRESSION MODELS

- Regression analysis performed between the variables like Total Sales, Price per volume, Price per unit (both with and without merch), Geographical Location, Form etc.,

- **Interaction Analysis:**

Interaction map is plotted for the various product prices and dollar sales of the Margarine product across various brands from 2018 to 2022.



REGRESSION MODELS

- Ordinary least square (OLS) regression analysis is conducted to examine the relationship between the variables Total sales and Price per unit and price per volume.

OLS Regression Results

Dep. Variable:	Total sales	R-squared (uncentered):	0.065
Model:	OLS	Adj. R-squared (uncentered):	0.065
Method:	Least Squares	F-statistic:	1315.
Date:	Tue, 09 May 2023	Prob (F-statistic):	0.00
Time:	04:17:27	Log-Likelihood:	-4.4744e+05
No. Observations:	37732	AIC:	8.949e+05
Df Residuals:	37730	BIC:	8.949e+05
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Price per Unit	4477.8626	120.944	37.024	0.000	4240.810	4714.916
Price per Volume	-579.7654	79.162	-7.324	0.000	-734.925	-424.606

Omnibus:	71039.004	Durbin-Watson:	2.022
Prob(Omnibus):	0.000	Jarque-Bera (JB):	161729682.129
Skew:	14.396	Prob(JB):	0.00
Kurtosis:	322.439	Cond. No.	3.24

Dependent variable(Y): Total sales
Independent variables(X): Price per Unit, Price per volume

F-statistic: 1315
P value: 0.0
R-squared: 0.065
Adjusted R-squared: 0.065

REGRESSION MODELS

- Regression model to understand the relation between total volume sold, price of product and the geographical location when no merch is provided.

R-squared: 0.10

- Regression model to understand the relation between the total volume sales with merchandise and without any merchandise based on the price per volume across various forms.

```
# provide a regression equation for predictions
intercept = reg_model.intercept_
coefficients = reg_model.coef_
```

```
print(intercept)
print(coefficients)
```

```
31671.18885126884
[-6892.09361255  3426.22805639  7482.9318352   3151.26130755
   818.43864286  3458.19398109  7699.82402937 48338.45943016
 -1020.05242531]
```

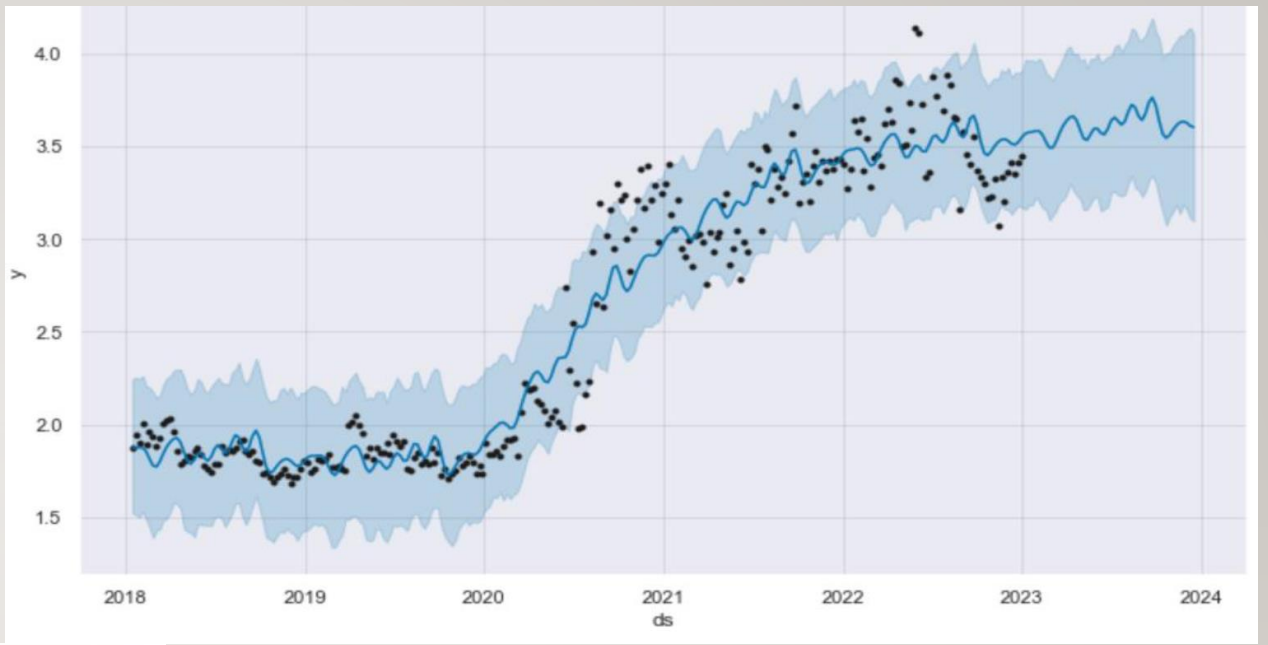
```
# provide a regression equation for predictions
intercept_merch = reg_model_merch.intercept_
coefficients_merch = reg_model_merch.coef_

print(intercept_merch)
print(coefficients_merch)
```

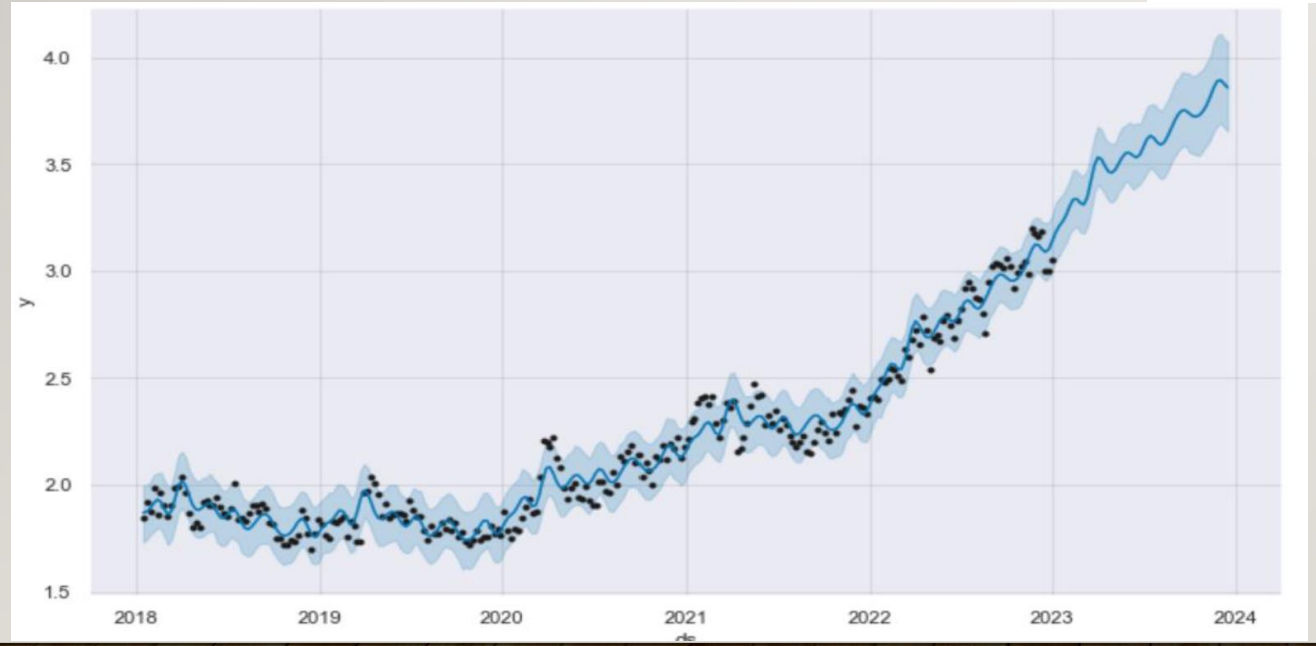
```
-11.390012000302335
[380.15317761 372.84851657 162.49948683 -12.62521698]
```

PREDICTIONS:

Weekly seasonality is interpreted, and the forecast is generated for the next 50 periods.

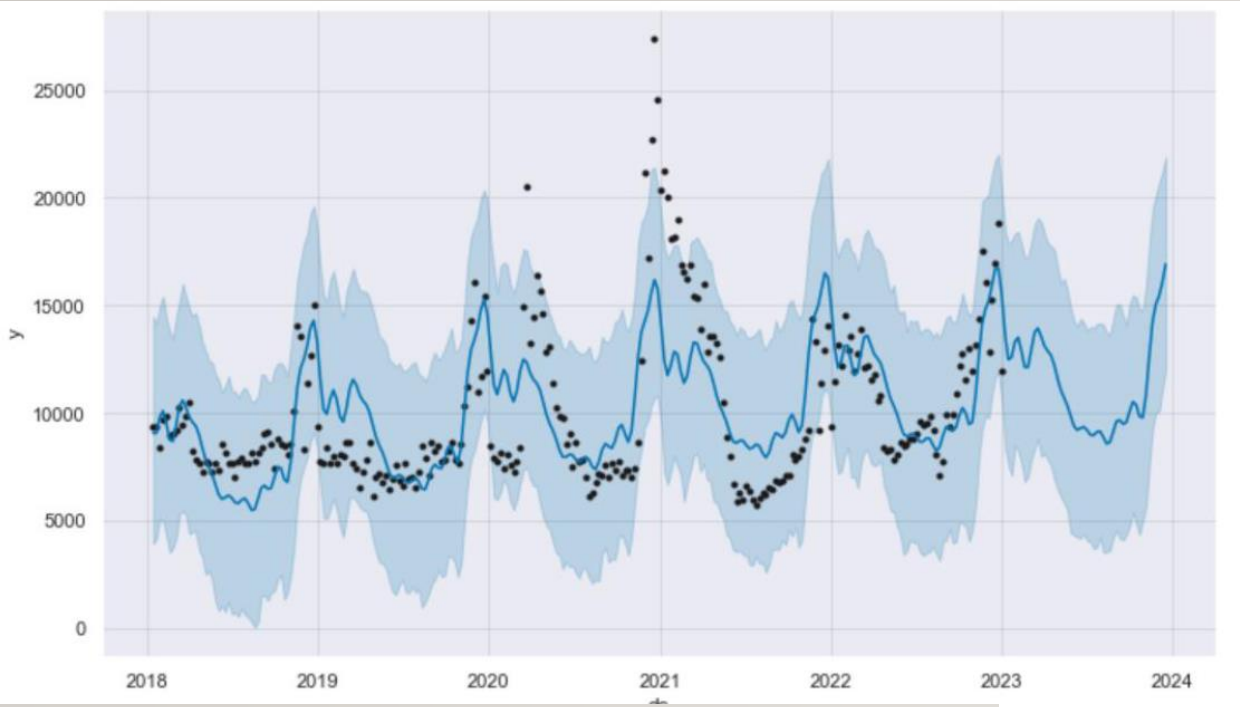
Forecast: **Price per Volume**





Forecast: **Price per Unit**

PREDICTIONS:



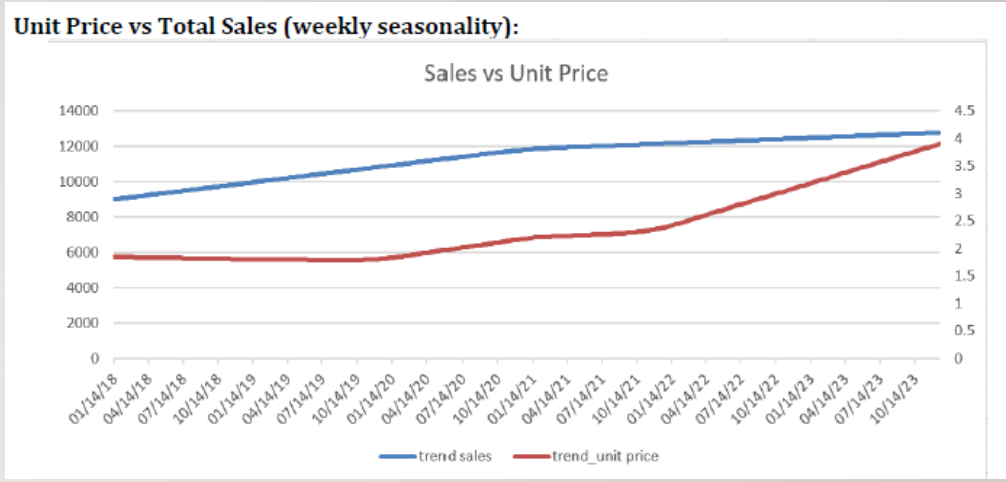
Forecast: **Total Sales**

Overall average trend for each month

PREDICTION OUTCOMES

- From the trend analysis, we can see that the total sales and Product price per unit values appear to be consistently increasing over time.
- The sales were increasing when the prices were moderate and when there is a hike in price per unit, the total sales tend to get saturated.

Unit Price vs Total Sales (weekly seasonality):



- Margarine experienced a steep price hike between 2020-2022 due to the pandemic, but total sales continued to increase gradually.
- The trend forecast indicates a decrease in sales despite a stable price per volume.

Volume Price vs Total Sales (weekly seasonality):

CONCLUSION:

- Conagra can increase sales by focusing on the final quarter of the year, improving sales by adding new customers, and providing healthy options to cater to the growing trend of healthy eating habits.

- Targeting the Asian community in some of the cities they reside in and optimizing prices in a profitable bracket can also increase sales and expand customer base.

- Strategies such as improving product quality, expanding the product line, enhancing customer experience, reducing costs, implementing a loyalty program, partnering with other businesses.

- By implementing a combination of these strategies, Conagra can not only increase sales and profits but also build customer loyalty and establish a strong brand identity in the market.