

# Group 16

# **Data-Driven Strategies to Unlock Future Growth Potential: A Predictive Marketing Analytics Approach**

**-by**

**Daniel Bond dab200000**

**Manoj Mareedu mxm220069**

**Yashasvi Pamu yxp220001**

**Nallam Paramkousam Nallam nxn220021**

**Pooja Reddy Donda pxd210043**

## Table of Contents:

Executive Summary.....	4
Introduction .....	5
Preprocessing of the data.....	6
Exploratory Data Analysis .....	6
Correlation Analysis .....	9
Hypothesis test.....	11
Hypothesis 1 .....	11
Hypothesis 2 .....	12
Hypothesis 3 .....	13
Anova .....	13
Regression Analysis .....	14
Linear Regression Model .....	14
Lasso Model .....	16
Polynomial Regression Model .....	17
Conclusion and Recommendations .....	19
References .....	21

# Executive Summary

Conagra is in need to identify insights that can help them optimize their portfolio of brands and accelerate growth. This report aims to analyze the sales data of table spreads to determine any trends in regionality, seasonality, brand effects, price elasticity, and possibly the impact of COVID-19. The study focused on three years, 2018 (pre-COVID), 2020 (during COVID), and 2022 (post-COVID). The analysis revealed a decrease in the brand count and an increase in mean Dollar Sales No Merch and Price per Unit No Merch over these three years. However, the Units Sold No Merch was not consistent, with a mean increase from 2018 to 2020 and a slight drop in 2022.

A linear regression model was developed to study the impact of various variables on the Units Sold No Merch without merchandising. The model included Base Volume Sales, Price per Unit No Merch, Price per Unit No Merch, ACV-weighted distribution No Merch, ACV-weighted distribution Any Merch, brands, regions, and years. The Units Sold No Merch were selected as the dependent variable, as they were considered a better measure of product demand. The analysis showed that the Southeast and Northeast regions had the best effect on sales, while the Plains region had the worst. Additionally, 2018 had the worst effect, while 2022 had the best. The brands had similar positive effects on sales, while regions and years had greater negative effects. The Price per Unit No Merch had a positive effect, while the Price per Unit Any Merch had a negative effect.

Hypothesis testing was conducted to determine if the differences in Units Sold No Merch between the years were significant enough for Conagra to not ignore. The analysis showed that the differences were significant, and an ANOVA test revealed that the Unit Sales No Merch between regions for each year were significantly different as well. A lasso test was conducted to identify the variables that had significant effects on Unit Sales No Merch, which included the Northeast region, BlueBonnet, Imperial RFG, Private Labels, and Smart Balance brands. A polynomial model was created using these significant variables to determine the combinative effects and provide Conagra with a tool to make further predictions.

In conclusion, the report recommends that Conagra should focus on studying the Northeast, Southeast, and Plains regions to understand what hidden variables make these regions successful or unsuccessful. If feasible, the study suggests that Conagra should consider not raising prices or lowering them slightly to gain customer goodwill and attract more customers during events like 2020. Finally, Conagra should prioritize the BlueBonnet brand over Smart Balance, as BlueBonnet had greater significance, while Smart Balance had a negative significance. Finally, the report recommends monitoring the Imperial RFG and Private Label brands, as they had significant effects

on Unit Sales No Merch and could be considered competitors to BlueBonnet. The linear regression equation used in the analysis was:

$$\text{Unit Sales No Merch} = (\text{Base Volume Sales}) + (\text{Price per Unit No Merch}) + (\text{Price per Unit Any Merch}) + (\text{ACV Weighted Distribution No Merch}) + (\text{ACV Weighted Distribution Any Merch}) + (\text{Brands}) + (\text{Regions}) + (\text{Years}).$$

## Introduction

Conagra Brands, a Fortune 500 American consumer packaged goods holding company, owns popular brands such as Birds Eye, Duncan Hines, Healthy Choice, and Marie Callender's. With over 18,000 employees and approximately \$11.054 billion in revenue this year, Conagra Brands aims to optimize their portfolio of brands and accelerate growth by finding insights with the tablespreads data. The company prides itself on combining data with culinary expertise to make great products. The business problem that Conagra Brands faces is the need to identify insights that can help them optimize their portfolio of brands and accelerate growth.

To address this problem, we will look to answer four questions: Is there regionality, seasonality, brand effect, or price elasticity within this data? By answering these questions, we hope to identify which regions and brands have a significant effect on the Unit Sales of Conagra Brands' products. Additionally, we hope to discover if there is a significant difference in sales between the three years (2018, 2020, 2022) and other variables such as regions or brands. Moreover, we will explore if any of the brands have a significant effect on the Unit Sales No Merch, and if the Price per Unit (No or Any Merch) affects Unit Sales No Merch, Regions, or Brands.

By identifying the regions that contribute the most to Unit Sales No Merch or brand performance, Conagra Brands will be able to adjust their strategies accordingly. For example, if we find that a particular region is significantly better than others in terms of contributing to Unit Sales No Merch, Conagra Brands can study what makes that region so beneficial and prioritize its focus on that region. They could also take what they learned about the best region and try applying it to other regions. Additionally, if we can show that a particular brand contributed more to the Unit Sales No Merch than others, Conagra Brands can prioritize its resources on that brand during times of need instead of brands that do not.

Furthermore, by analyzing the impact of prices on sales, Conagra Brands can identify which regions or brands it might want to lower or raise prices for. This information can be used to determine which brands or regions are more reliable during future crises and help the company adjust its pricing strategies accordingly. Ultimately, by answering these questions, Conagra Brands

will be able to optimize their portfolio of brands and accelerate growth by making data-driven decisions on where they should optimize and which brands.

## **Preprocessing of the data**

Preprocessing of data is a crucial step in data analysis that involves modifying and organizing datasets to extract meaningful insights and conclusions. In our study, we followed a rigorous process that involved replacing null values with zeros and removing duplicate entries to ensure the accuracy of our data. We did not want to remove the null observations because it could skew the data. We then extracted the year and brand name from existing columns and created separate columns for each. This allowed us to conduct further analysis by adding dummy variables for brands, years, and geography.

To examine the potential impact of the COVID-19 pandemic on sales, we merged datasets from three specific years, namely 2018, 2020, and 2022. By selecting these years, we aimed to capture and control the pre-crisis, crisis, and post-crisis or recovery periods' pattern for these three situations. This allowed us to observe how the pandemic affected sales during different stages of the pandemic.

To maintain the relevance of our analysis, we excluded the years 2019 and 2021 from our study. These years were affected differently by COVID-19 lockdowns in various states and including them in our analysis would have interfered with our research questions. By excluding these years, we were able to focus on the three most relevant years for our study and ensure that our findings were accurate and reliable.

Overall, the process of modifying and organizing datasets is critical to ensure the accuracy and reliability of data analysis. By carefully selecting and preparing the data, we can obtain meaningful insights and draw accurate conclusions, which are essential for making informed decisions and taking appropriate actions.

## **Exploratory Data Analysis**

For the tablespreads data, there are 24 different variables and about 270000 observations with 5000 to 7000 nulls for each year's data set. Below are the different variables.

Geography	Qualitative
Time	Qualitative
Product Description	Qualitative
UPC 13 digit	Quantitative
Dollar sales no Merch	Quantitative
Dollar sales any merch	Quantitative
Unit sales no merch	Quantitative
Unit sales any merch	Quantitative
Volume sales no merch	Quantitative
Volume sales any merch	Quantitative
Price per unit	Quantitative
Price per unit no merch	Quantitative
Price per unit any merch	Quantitative
Price per volume	Quantitative
Price per volume no merch	Quantitative
Price per volume any merch	Quantitative
Acv weighted distribution no merch	Quantitative
Acv weighted distribution any merch	Quantitative
Base unit sale	Quantitative
Base volume sale	Quantitative
Base dollar sales	Quantitative
Incremental unit	Quantitative
Incremental Volume	Quantitative
Incremental dollar	Quantitative

The data sets show the table spread products and different statistics such as how much it sold in dollars, units sold, any merchandising done, geography, etc. for each ending week of the year.

## 2018

```
df2018[['Dollar Sales No Merch', 'Unit Sales No Merch', 'Price per Unit No Merch']].describe()
```

	Dollar Sales No Merch	Unit Sales No Merch	Price per Unit No Merch
count	2.706520e+05	2.706520e+05	270652.000000
mean	2.092813e+04	6.556351e+03	3.941502
std	8.808144e+04	3.381238e+04	3.087638
min	0.000000e+00	0.000000e+00	0.000000
25%	1.936852e+02	5.237671e+01	2.490000
50%	1.583725e+03	4.794341e+02	3.518908
75%	1.001506e+04	2.922098e+03	4.873482
max	4.218969e+06	2.182544e+06	82.400000

## 2020

```
[37] df2020[['Dollar Sales No Merch', 'Unit Sales No Merch', 'Price per Unit No Merch']].describe()
```

	Dollar Sales No Merch	Unit Sales No Merch	Price per Unit No Merch
count	2.683420e+05	2.683420e+05	268342.000000
mean	2.548426e+04	7.752487e+03	4.024998
std	1.163257e+05	4.193194e+04	2.750889
min	0.000000e+00	0.000000e+00	0.000000
25%	2.008605e+02	5.221596e+01	2.682149
50%	1.655651e+03	4.770075e+02	3.616332
75%	1.120619e+04	3.211990e+03	4.942362
max	5.963802e+06	2.519289e+06	199.990000

## 2022

```
[38] df2022[['Dollar Sales No Merch', 'Unit Sales No Merch', 'Price per Unit No Merch']].describe()
```

	Dollar Sales No Merch	Unit Sales No Merch	Price per Unit No Merch
count	2.472100e+05	2.472100e+05	247210.000000
mean	3.006908e+04	7.319866e+03	4.676040
std	1.335760e+05	3.862062e+04	2.486688
min	0.000000e+00	0.000000e+00	0.000000
25%	2.555459e+02	5.740887e+01	3.228211
50%	1.950369e+03	4.640000e+02	4.260518
75%	1.301485e+04	3.137044e+03	5.708953
max	7.427815e+06	2.645258e+06	199.990000

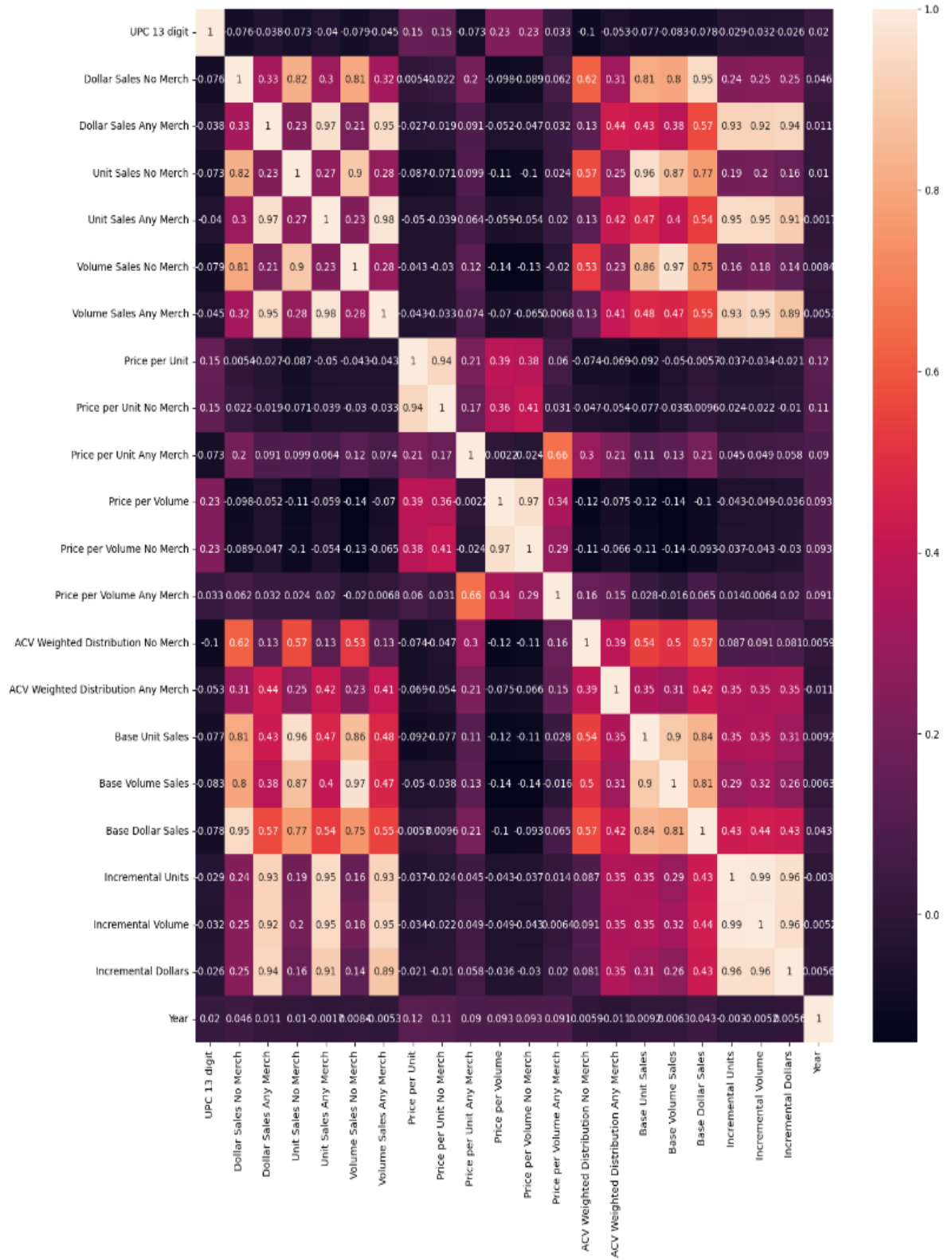


In order to gain a preliminary understanding of the dataset, we examined the summary statistics of the 24 variables. Our analysis revealed two main patterns worth noting: the counts and means of the variables. Specifically, we observed that the count of observations decreased from 2018 to 2020, possibly due to factors such as brand discontinuations, supply chain issues, or other reasons related to the COVID-19 pandemic. The count then decreased again from 2020 to 2022, suggesting that even more brands were affected by the pandemic and were dropped from the dataset.

In terms of the means, we found that the Dollar Sales No Merch and Price per Unit No Merch increased from 2018 to 2020 and continued to rise from 2020 to 2022. This indicates that there was an increasing demand for these products throughout the three years, and customers were willing to pay higher prices despite the increased cost. However, the trend for Unit Sales No Merch was not consistent with this pattern. While it increased from 2018 to 2020, it dropped below 2020 levels in 2022 but remained above 2018 levels. This could be attributed to the COVID-19 lockdowns, during which people had to cook at home and therefore purchased more groceries. In 2022, the decreased Unit Sales No Merch could be a result of the rising prices leading customers to reduce their purchases, or due to the lifting of lockdown restrictions allowing people to eat out more frequently. Overall, our exploratory data analysis highlights some interesting patterns in the dataset that can be further investigated in future analyses.

## **Correlation Analysis**

The correlation analysis showed interesting insights for the table spreads market. Specifically, the variables related to No Merch and Any Merch showed strong correlations with themselves, but weak correlations with Price per Unit variables. On the other hand, No Merch Dollar Sales was highly correlated with Base Unit and Dollar Sales, while Any Merch showed high correlation with Incremental variables. Notably, No merch and Any Merch were correlated with each other in cooking oils but not in table spreads. This suggests that merchandising products may be boosting sales of certain products that complement the non-merchandising products for cooking sprays at least. In contrast, the different regions did not show any significant correlations with other variables. Surprisingly, Price per Unit/Volume for both Any Merch and No Merch were not highly correlated with any other variables, including Sales or Units Sold. These insights will be used to build a model that will be beneficial to Conagra while not suffering from multicollinearity.



Heat Plot of Correlation Matrix

### **Final Formula:**

Unit sales no merch = (Base Volume Sales) + (Price per Unit No Merch) + (Price per Unit Any Merch) + (ACV Weighted Distribution No Merch) + (ACV Weighted Distribution Any Merch) + (brands) + (regions) + (years).

To ensure that our dependent variable was not influenced by external factors, we chose Unit Sales No Merch as our dependent variable, as opposed to dollar sales. However, we found that only Base Volume Sales had a correlation with Unit Sales No Merch and was necessary to provide an explanatory variable. To investigate the variation in sales in different regions during the years 2018, 2020, and 2022, we used the "regions" variable. To evaluate the correlation of different brands with Unit Sales No Merch, we created a dummy variable called Brands. Additionally, we created dummy variables for 2018, 2020, and 2022 to see the differences in Unit Sales No Merch for each year. We also found that ACV Weighted Distribution, which measures the amount of product shipped to stores versus the store's size, helped us figure out how many units a store was moving and how much was sitting in storage. Although Base Volume Sales is correlated with Unit Sales No Merch, it is a calculated variable and was found to be the most accurate independent variable for our model. Overall, our approach was effective in minimizing the influence of external factors on our dependent variable and allowed us to identify the key independent variables that explain Unit Sales No Merch.

## **Hypothesis test**

In order to better understand the significance of changes in Unit Sales No Merch across different years and regions, we conducted a series of hypothesis tests. These tests will provide valuable insights for Conagra, helping them to determine whether the observed changes in Unit Sales No Merch are meaningful or if they can be dismissed as random fluctuations.

### **Hypothesis 1**

*H0*: No significant difference between Unit Sales No Merch between 2018 and 2020 data.

*Ha*: Significant difference

### **Result:**

t-statistic 2018-2020: -11.522 p-value 2018-2020: 0.00

**Inference:**

We conducted a statistical test to examine the significance of the changes in Unit Sales No Merch between the years 2018 and 2020. The null hypothesis ( $H_0$ ) assumed that any difference in Unit Sales No Merch between the two years was due to chance, while the alternative hypothesis ( $H_a$ ) posited that there was a significant difference between the two years' Unit Sales No Merch, potentially due to external factors such as the COVID-19 pandemic.

Our test yielded a t-statistic of -11.522 and a p-value of 0.00, which indicates a very low probability of observing such a significant test statistic under the null hypothesis. Therefore, we reject the null hypothesis and accept the alternative hypothesis that the difference in Unit Sales No Merch between the two years is meaningful and not due to chance. This finding suggests that Conagra should pay attention to the change in Unit Sales No Merch and consider any external factors that may have contributed to it.

**Hypothesis 2**

$H_0$ : No significant difference between Unit Sales No Merch between 2020 and 2022 data.

$H_a$ : Significant difference

**Result:**

t-statistic 2020-2022: 3.856 p-value 2020-2022: 0.000

**Inference:**

In order to test the hypothesis that there was a significant difference in Unit sales No Merch between the years 2020 and 2022, we conducted a t-test using the t-statistic. Our results showed that the p-value was 0.000, which is smaller than the commonly used significance level of 0.05. Therefore, we can conclude that the difference in Unit Sales No Merch between these two years is statistically significant, and we reject the null hypothesis in favor of the alternative hypothesis.

However, it is important to note that other factors, such as the COVID-19 pandemic and changes in consumer behavior, may have influenced these sales. Therefore, it is essential to consider these factors when interpreting the results.

Overall, our findings suggest that the difference in Unit Sales No Merch between 2020 and 2022 should not be disregarded by Conagra. These results could potentially inform future business decisions and strategies.

### Hypothesis 3

$H_0$ : No significant difference between Unit Sales No Merch between 2018 and 2022 data.

$H_a$ : Significant difference

### Result:

t-statistic 2018-2022: -7.539 p-value 2018-2022: 0.000

### Inference:

We conducted a hypothesis test to determine if there was a significant difference in Unit Sales No Merch between the years 2018 and 2022. Our null hypothesis suggested that there was no difference, while the alternative hypothesis indicated a significant difference. The p-value of 0.000 was obtained from the test, indicating strong evidence against the null hypothesis. This result suggests that the difference in Unit Sales No Merch between 2018 and 2022 is significant and not due to chance. Therefore, Conagra should consider this difference when making business decisions.

Overall, the results of all three hypotheses indicate that Unit Sales No Merch are consistently significant and should not be disregarded, despite not following the same pattern as Price per Unit and Dollar Sales No Merch. Furthermore, the high T-statistic for the change between 2018 and 2020 suggests that this change was particularly significant.

### Anova

#### 2018 Results:

----- ANOVA test 2018 -----				
	sum_sq	df	F	PR(>F)
C(Geography)	4.200045e+12	8.0	465.516244	0.0
Residual	3.052291e+14	270643.0	NaN	NaN

## 2020 Results:

----- ANOVA test 2020 -----				
	sum_sq	df	F	PR(>F)
C(Geography)	6.378447e+12	8.0	459.656416	0.0
Residual	4.654422e+14	268333.0	NaN	NaN

## 2022 Results:

----- ANOVA test 2022 -----				
	sum_sq	df	F	PR(>F)
C(Geography)	5.644077e+12	8.0	480.34098	0.0
Residual	3.630810e+14	247201.0	NaN	NaN

Furthermore, we conducted an ANOVA test to evaluate whether there are noteworthy variations in Unit Sales No Merch across different regions. It is known that different regions have diverse demographics and other variables that may affect their sales, but it was important to identify whether these differences were significant enough to require further investigation by Conagra. The results of the P-value analysis demonstrated that there are indeed significant differences in Unit Sales No Merch between the regions for each of the years. Conagra should pay close attention to these variations and explore the underlying factors that contribute to these differences. This investigation could include examining the unique practices of specific regions, demographic variances, or other factors that might impact consumer behavior. By doing so, Conagra may discover patterns that can be replicated in other regions to increase their sales and enhance their overall performance.

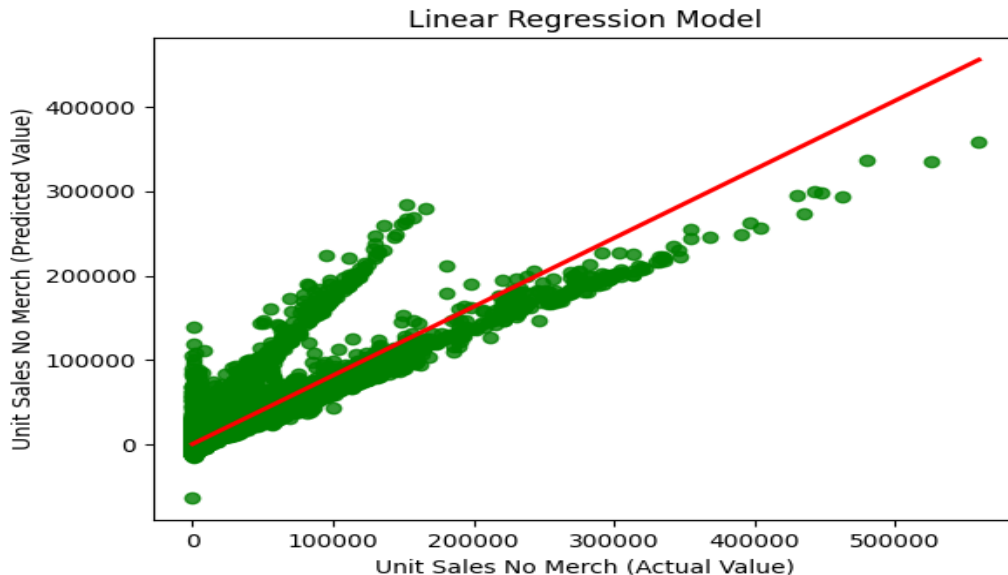
# Regression Analysis

## Linear Regression Model

Our team constructed a linear regression model by combining the data from three years and using the formula mentioned earlier. This enabled us to observe the impact of various independent variables on the dependent variable, which is Unit Sales No Merch. The coefficients obtained from the model revealed some interesting findings. We found that the Price per Unit No Merch products had a positive effect on the unit sales, while the Price per Unit Merch products

had a negative effect. The weight distribution for both Merch and No Merch products had a similar effect but was not as great as the Price per Unit. The Base Volume Sales, which was highly correlated with Unit Sales No Merch, did not have a significant effect on the unit sales.

Furthermore, all regions had a negative effect on unit sales, but the Southeast had the least negative impact, followed by the Northeast, while the Plains region had the most negative impact. This indicated that products were more likely to sell better in the North and Southeast than in the Plains. Brands played a crucial role in driving unit sales, and all brands had a positive effect of similar magnitude. However, the effect of brands was less significant than the effect of regions. In terms of the time variable, the worst year was 2018, and the least worst was 2022, although all years had a negative effect. This suggests that people were more likely to buy fewer products in 2018, possibly due to the wider range of food choices available. It is worth noting that we had additional data beyond the three years due to a few observations that spilled over to the next year.



*Linear regression model plot: Unit Sales No Merch (Actual Value vs Predicted Value)*

## **Lasso Model**

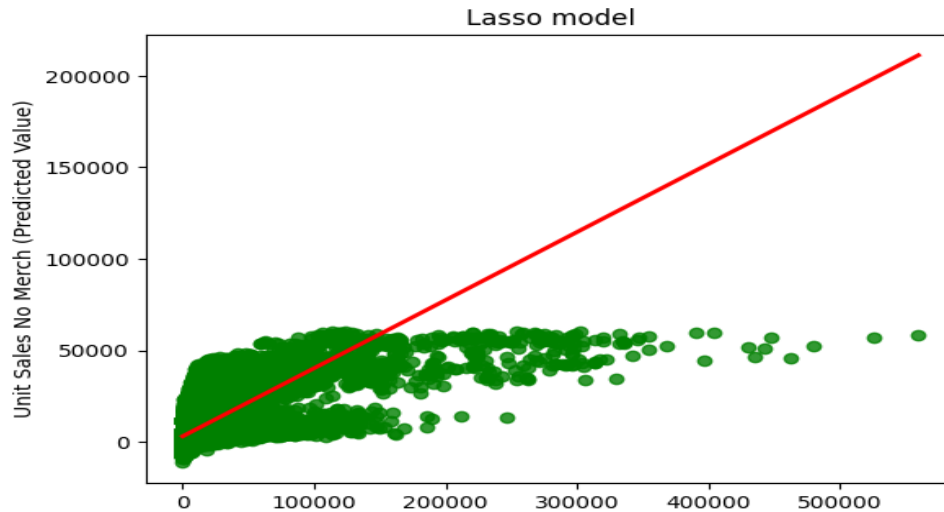
To better understand the factors influencing Unit Sales No Merch, a linear regression was conducted initially to determine the coefficients of each variable in relation to the target variable. However, this analysis was limited as it did not account for the significance of each variable. To address this, a Lasso model was employed to determine which variables were worth investigating further. Following the Lasso analysis, it was found that the Price per Unit Any Merch was still negatively significant, while the ACV weighted distribution proved to be more significant than the Price per Unit values. Interestingly, the No Merch ACV had a larger magnitude than its Merch counterpart, indicating that stores carrying more non-merchandise products may have higher Unit sales.

Regionally, California had a positive significance, while the Northeast was the most positively significant region. Surprisingly, the Southeast, which performed well in the linear regression model, did not appear significant in the Lasso model. The Plains region had the most significant negative magnitude, making it the least favorable region for Unit Sales No Merch in both models. Of the brands, only a few proved to be significant. Imperial RFG had the highest significance, followed by BlueBonnet, private labels, and Smart Balance. Conagra should focus on Imperial RFG and private labels, as they could be competitors to Conagra's own BlueBonnet label. Additionally, further analysis is needed to determine why Smart Balance has a negative coefficient.

None of the years were found to be significant, except for 2018, which had a negative effect on Unit sales No Merch.

Overall, the Lasso model provided valuable insights into which variables, brands, and regions were most significant for Unit Sales No Merch, and which were not. By focusing on the significant variables, brands, and regions, Conagra can make informed decisions to improve their sales strategy and optimize profits.



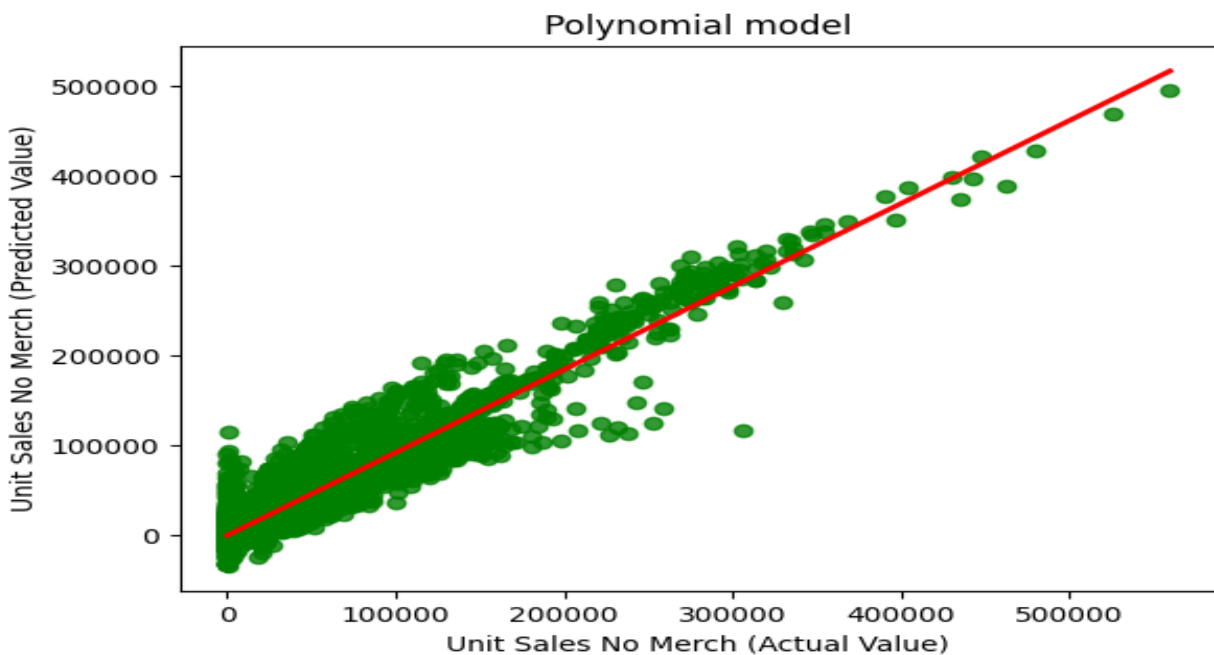
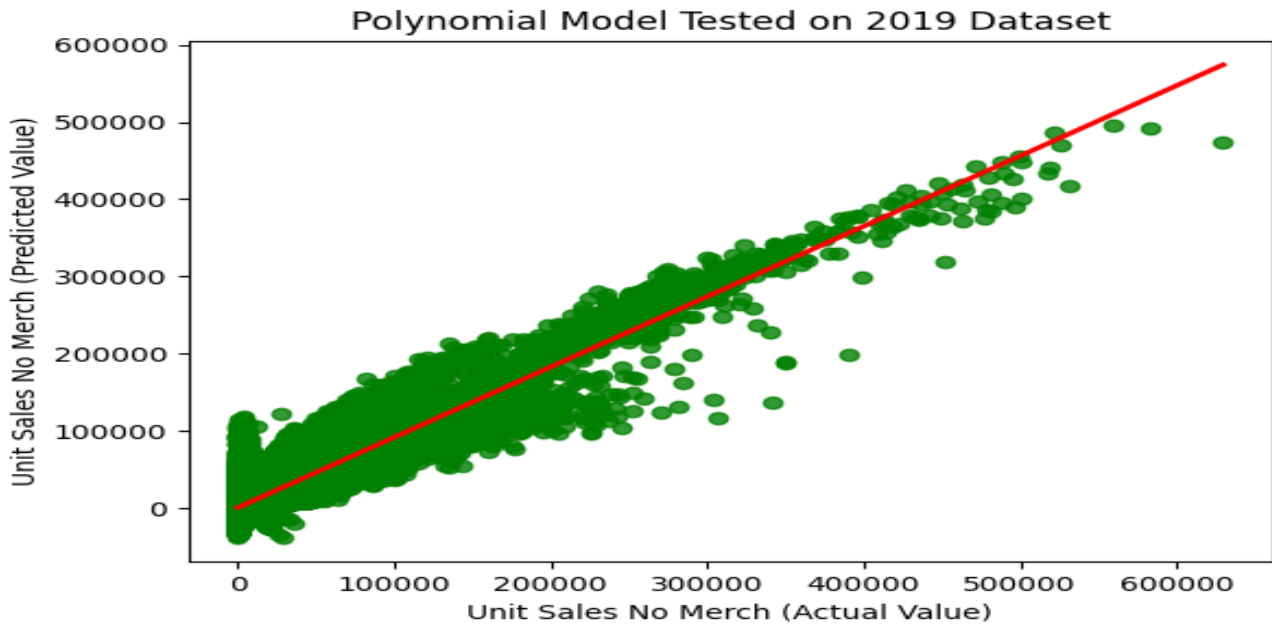


*Lasso model plot: Unit Sales No Merch (Actual Value vs Predicted value)*

## Polynomial Regression Model

We developed a polynomial model to further investigate the significant variables identified by the Lasso model. This allowed us to assess the combined effects of these variables on Unit Sales No Merch, providing valuable insights into future sales trends for Conagra. Our model included Base Volume Sales, which was found to improve the model's accuracy to approximately 81% and was significant in the Lasso model. We included only the most significant brands identified by the Lasso model, namely Smart Balance, Private Label, Imperial RFG, and BlueBonnet.

To test the accuracy of the polynomial model, we used the year 2019 data as a testing year, and the results were encouraging. This suggests that our polynomial model can be a reliable tool for forecasting future sales trends and can assist in making informed business decisions by identifying the key factors influencing Unit Sales No Merch.



*Plots of Polynomial model Actual Values vs Predicted value.*

In summary, the polynomial model is a valuable tool for gaining a more in-depth understanding of the complex relationships between various variables and their impact on Unit Sales No Merch. With this model, Conagra can make informed decisions about which variables to prioritize to enhance sales and profitability. By identifying the most influential independent variables, Conagra can focus their efforts on improving those variables to increase Unit sales No

Merch. Additionally, the model can help identify any non-linear relationships between variables that may have been overlooked with a simple linear regression model. Overall, utilizing the polynomial model can assist Conagra in making data-driven predictions that involve the patterns we witnessed in the other models.

## **Conclusion and Recommendations**

### **Regionality:**

Based on our analysis, we recommend that Conagra direct their attention to three regions which are Northeast, Southeast, and Plains. The Northeast was found to be the most significant and positively impactful region in the Lasso model, while the Southeast, although not significant in the Lasso model, had a larger effect in the linear model. On the other hand, the Plains region exhibited the most negative impact in both models, with the largest magnitude in the Lasso model. By focusing on these three regions and identifying the underlying factors that contribute to their success or failure, Conagra may be able to replicate these effects in other regions by implementing similar strategies. This could involve examining factors such as product placement and supply chain efficiency and adjusting their strategies accordingly. Moreover, if Conagra can identify the reasons for success in these regions, they could apply these insights to improve performance in the Plains region and potentially other regions as well.

### **Seasonality:**

Based on our analysis of the summary data and models, it appears that the COVID-19 pandemic had a positive impact on this particular business sector. The statistics indicated that although the Price per Unit No Merch increased, the Dollar Sales also went up. The Unit Sales No Merch increased from 2018 to 2020 and decreased slightly in 2022, but still remained higher than in 2018. One theory is that during the pandemic lockdowns, people had to rely on cooking at home, leading to an increase in demand for food products, which subsequently caused the rise in dollar and Unit Sales No Merch in 2020. However, this also led to a rise in prices due to supply chain issues and increased demand. In 2022, even though the lockdowns were lifted, supply chain issues persisted, causing continued price inflation. This might have led to some people being more selective in their purchases, resulting in a decrease in Unit Sales No Merch. Nonetheless, people still preferred cooking at home due to cost-saving measures, which could explain why the Unit Sales No Merch were still above the 2018 levels. Our analysis also revealed that the year 2018 had a negative impact, possibly indicating a lower demand for food products during that year. The region played a more significant role than the year, and the magnitude of the years seemed to be similar, with 2018 being the worst and 2022 being the best.

## Brands:

Upon analyzing the results of linear regression, we observed that all the brands showed a similar positive effect on Units Sold No Merch, but their individual impact was relatively smaller than that of regions and years. However, when we applied the lasso method, only four brands - BlueBonnet, Imperial RFG, private labels, and Smart Balance - were identified as significant. Among these, Imperial RFG stood out with double the significance of BlueBonnet. It would be prudent for Conagra to investigate Imperial RFG further and identify the variables that contribute to its success. The company should also explore why Smart Balance is showing negative significance and determine whether additional merchandising efforts may be needed or if it may be appropriate to phase it out gradually. BlueBonnet, on the other hand, showed significant performance, making it an important brand to consider for future investment. Overall, these findings provide valuable insights for Conagra to optimize its product portfolio and allocate resources more effectively.

## Price:

As previously mentioned, the summary statistics indicate a consistent increase in the Price per Unit of No Merch products from 2018 to 2022. Interestingly, our linear model revealed that the impact of Price per Unit No Merch was greater than its counterpart, Any Merch. However, the lasso model yielded unexpected results, as Price per Unit No Merch was not found to be significant, while Price per Unit Any Merch showed a negative significance with respect to Units Sold No Merch. These findings suggest that during times of high demand for food, customers may not be as sensitive to price when making purchasing decisions. Conagra could leverage this insight by maintaining consistent prices, which may help earn customer loyalty and trust. Furthermore, it may help Conagra outperform competitors who choose to raise prices during times of need. The negative significance of Any Merch on Unit Sales No Merch also suggests that Conagra could benefit from reducing its merchandising efforts for Unit Sales No Merch.

# References

1. Appendix:

<https://docs.google.com/document/d/1LeK6n0s61gncXDBWjvW6CRCEgDHy9i9ztIEDHXdq7eg/edit?usp=sharing>

2. Tablespreads Category Overview:

[https://www.dropbox.com/sh/ux0lvv0nckh17nx/AAAUiOIqcA9TfO11V3lpFCIYa/Conagra%20Data%20Files?dl=0&preview=UTD+Final+Group+Project.docx&subfolder\\_nav\\_tracking=1](https://www.dropbox.com/sh/ux0lvv0nckh17nx/AAAUiOIqcA9TfO11V3lpFCIYa/Conagra%20Data%20Files?dl=0&preview=UTD+Final+Group+Project.docx&subfolder_nav_tracking=1)

3. Pandas Data frame:

[pandas.DataFrame — pandas 2.0.1 documentation \(pydata.org\)](#)

4. Lasso Model:

[sklearn.linear\\_model.Lasso — scikit-learn 1.2.2 documentation](#)

5. Polynomial Regression:

[sklearn.preprocessing.PolynomialFeatures — scikit-learn 1.2.2 documentation](#)

6. Python Code:

[https://drive.google.com/file/d/1U9ubznsQeRmRHKaNwvOtdofv94bMrC4X/view?usp=share\\_link](https://drive.google.com/file/d/1U9ubznsQeRmRHKaNwvOtdofv94bMrC4X/view?usp=share_link)