

## **Final Project**

### **Tableau**

(Sripal Reddy Nomula, UID: 2001093288)

The given data contains information related to sales and promotional activities of products in four categories: mouthwash, pretzels, frozen pizza, and boxed cereal. The data spans a period of 156 weeks, starting from January 2009 and ending in December 2011. The dataset comprises various variables that provide detailed information about the stores, products, sales, and promotions. Let's go through the descriptions of the variables:

**ADDRESS\_CITY\_NAME:** This variable is a part of the "store lookup" table and represents the city in which the store is located.

**ADDRESS\_STATE\_PROV\_CODE:** Also part of the "store lookup" table, this variable denotes the state or province in which the store is situated.

**AVG\_WEEKLY\_BASKETS:** Another variable from the "store lookup" table, it indicates the average number of baskets (i.e., purchases) sold in the store on a weekly basis.

**BASE\_PRICE:** This variable, present in the "data" table, represents the base price of the product. The base price is the initial price of the product before any discounts or promotions are applied.

**MANUFACTURER:** Part of the "products lookup" table, this variable indicates the manufacturer or company that produces the product.

**CATEGORY:** Also a part of the "products lookup" table, this variable specifies the category to which the product belongs, such as mouthwash, pretzels, frozen pizza, or boxed cereal.

**DESCRIPTION:** Found in the "products lookup" table, this variable provides a description or name of the product.

**DISPLAY:** This variable, present in the "data" table, indicates whether the product was a part of an in-store promotional display. An in-store display is a marketing tactic where products are showcased prominently within the store to attract customers.

**FEATURE:** Another variable from the "data" table, it denotes whether the product was featured in an in-store circular. In-store circulars are promotional materials that advertise products and discounts.

**HHS:** This variable, present in the "data" table, represents the number of purchasing households. It indicates the number of unique households that purchased the product.

**MSA\_CODE:** This variable is a part of the "store lookup" table and stands for Metropolitan Statistical Area code. It represents a geographic region with a high core population density and close economic ties throughout the surrounding areas.

**PARKING\_SPACE\_QTY:** Another variable from the "store lookup" table, it provides the number of parking spaces in the store parking lot.

**PRICE:** Present in the "data" table, this variable indicates the actual amount charged for the product at the shelf, which may differ from the base price if there are promotions or discounts.

**WEEK\_END\_DATE:** This variable in the "data" table represents the date at the end of the week to which the sales and promotional information pertains.

**SALES\_AREA\_SIZE\_NUM:** Part of the "store lookup" table, this variable specifies the square footage of the store, indicating its size.

**STORE\_APPEAL:** Also from the "store lookup" table, this variable represents the retailer's designated store appeal, which may refer to the overall attractiveness or ambiance of the store.

**SPEND:** This variable in the "data" table denotes the total spend (i.e., total sales) generated by the product, measured in dollars.

**STORE\_NUM:** This variable is present in both the "data" table and the "store lookup" table, and it serves as a unique identifier for each store.

**SUB\_CATEGORY:** Part of the "products lookup" table, this variable provides a sub-category classification for the product.

**TPR\_ONLY:** Found in the "data" table, this variable indicates whether the product had a temporary price reduction only, meaning it was reduced in price but not part of any in-store display or advertisement.

**UNITS:** This variable in the "data" table represents the number of units (i.e., individual products) sold.

**UPC:** This variable is present in both the "data" table and the "products lookup" table, and it stands for Universal Product Code. It is a specific identifier for each product.

**VISITS:** Another variable from the "data" table, it represents the number of unique purchases (baskets) that included the product.

**PRODUCT\_SIZE:** Part of the "products lookup" table, this variable provides information about the package size or quantity of the product.

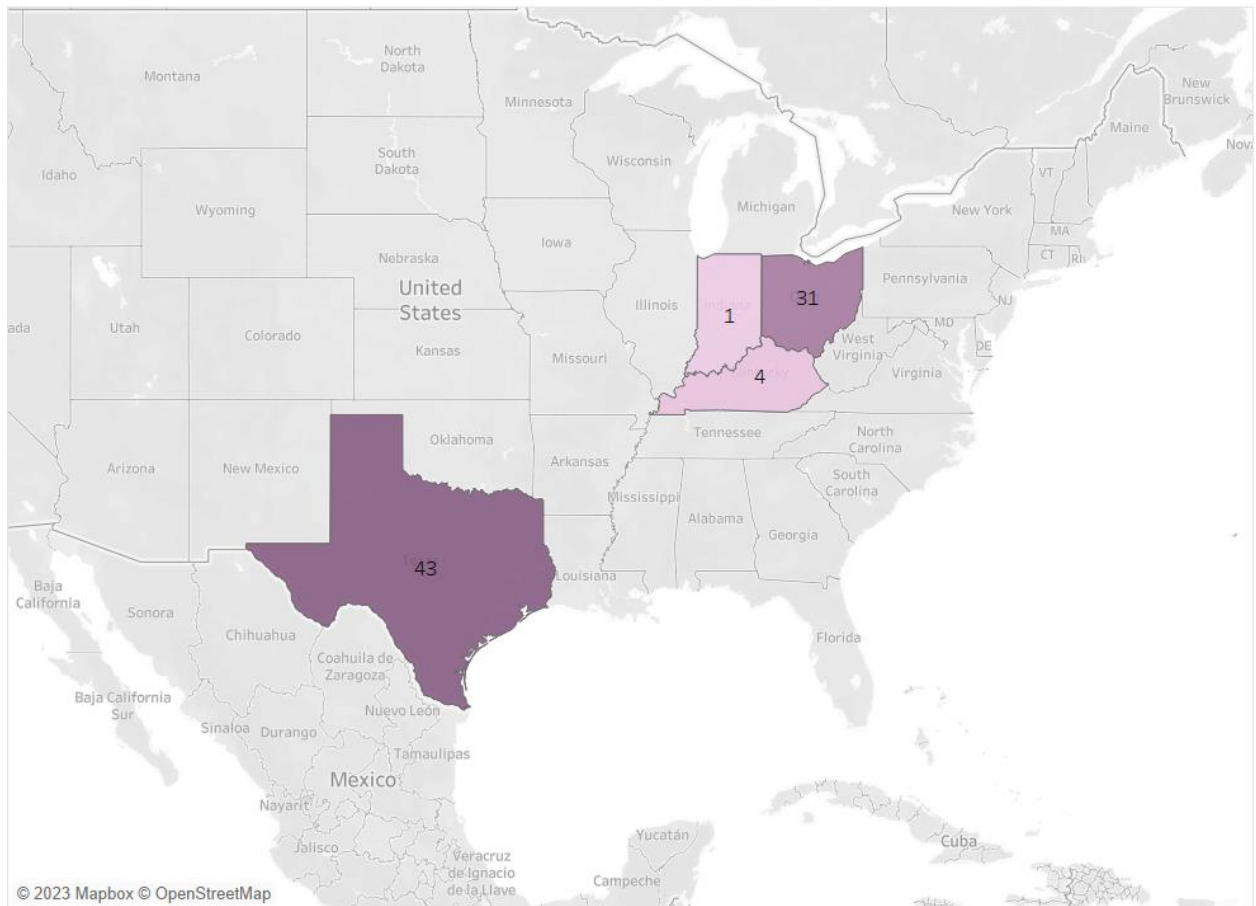
**Data source and Integration:** We have two files: "College Life Data Challenge" and "MSAData Text File." The Excel file contains three sheets that need to be integrated to extract relationships smoothly. The steps followed for integration are as follows:

1. Connect the "MSA" and "Store Lookup" tables based on the "MSA-Code" column.
2. Connect the "Store Lookup" and "Transaction Data" tables based on the "StoreID" column.
3. Connect the "Product Lookup" and "Transaction Data" tables based on the "UPC" column.

## Worksheets and Visualizations of EDA and Analysis of various trends:

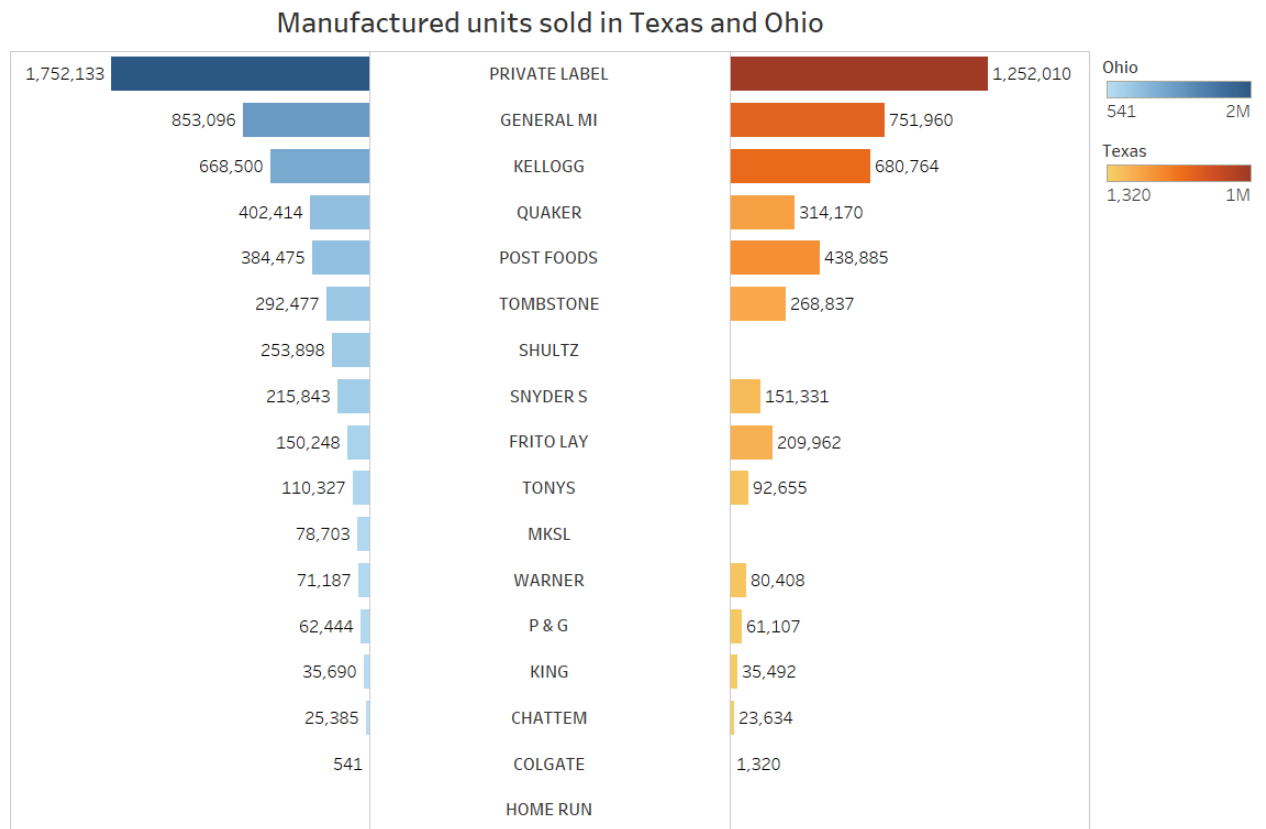
### 1. Stores in various Geographical Locations

Stores in diferent Geographic locations



The geographic distribution of our stores is evident from the data, revealing that the majority of our stores are concentrated in four states: Texas, Ohio, Kentucky, and Indiana. Texas stands out as the state with the highest number of stores, accounting for 41 locations. Following closely is Ohio with 31 stores, while Kentucky has 4 stores, and Indiana has a single store. This distribution highlights the regional focus of our business, with a significant presence in Texas and expanding into other states to cater to customer demand and capture new markets.

## 2. Units sold in Texas and Ohio by various manufacturers



Sum of Ohio, sum of Zeroline and sum of Texas for each Manufacturer. For pane Sum of Ohio: Color shows sum of Ohio. The marks are labeled by sum of Ohio. For pane Sum of Texas: Color shows sum of Texas. The marks are labeled by sum of Texas.

The above plot highlights the disparity in the number of stores between Texas (TX) and Ohio (OH) compared to Indiana (IN) and Kentucky (KY). The **butterfly plot**, on the other hand, focuses on the units sold by various manufacturers in TX and OH, shedding light on their competition and areas for improvement. Notably, Private Label emerges as the leader in units sold, with approximately 3 million units in both states, followed by General MI and Kellogg. One intriguing observation is that Shultz and MSTL have minimal presence in Ohio, despite having sales in Texas.

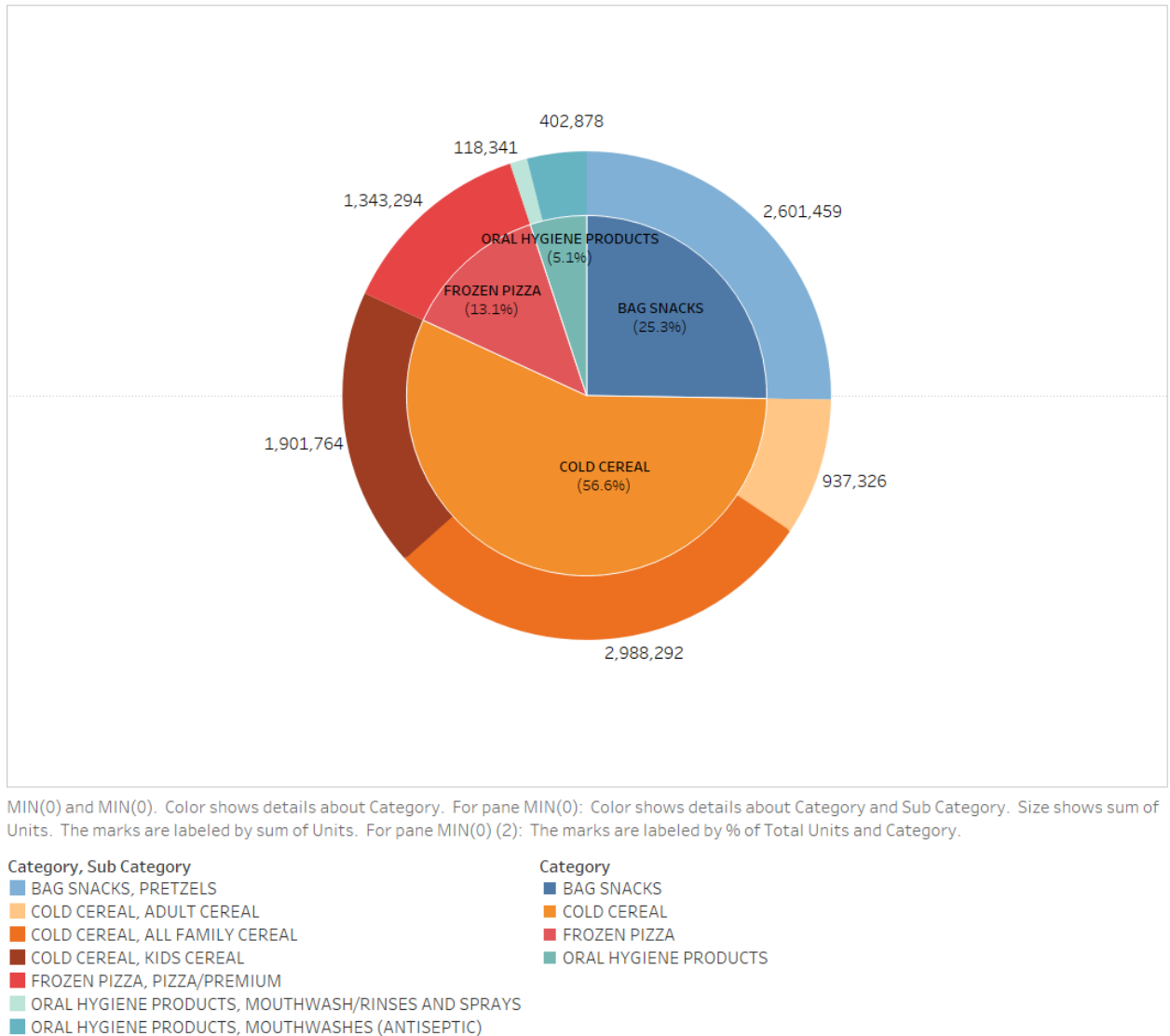
The butterfly plot serves as a valuable tool to discern market dynamics and opportunities for growth. It reveals the dominance of Private Label in both TX and OH, signaling their successful strategies and strong appeal to consumers in these regions. For manufacturers like General MI and Kellogg, being in the top ranks indicates their competitive performance.

However, the contrast between the presence of Shultz and MSTL in Texas and their absence in Ohio is intriguing. This points to potential challenges or untapped opportunities in the Ohio market for these manufacturers. Further analysis is essential to understand the reasons behind this discrepancy and to devise strategies to expand their presence in Ohio.

The data offers valuable insights for manufacturers to leverage their strengths and address weaknesses in specific regions. It underscores the significance of tailoring marketing and distribution strategies based on regional preferences and competition to thrive in diverse markets like Texas and Ohio.

### 3. Revenue/Sales generated by category and Sub- category.

Sales by category and sub-category



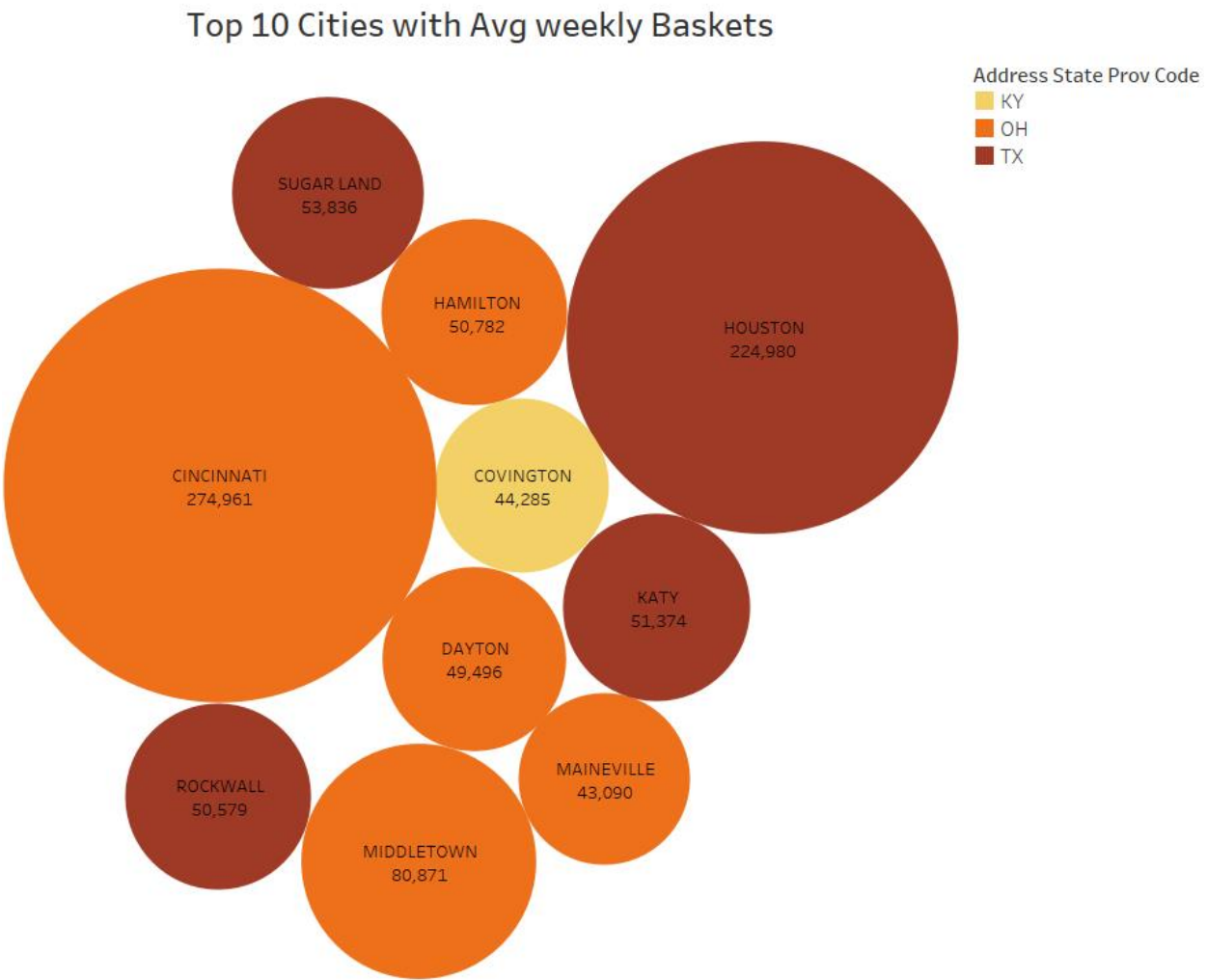
The Sunburst chart offers valuable insights into the distribution of sales and revenue among different categories and sub-categories. In the first inner pie chart, it is evident that cold-cereal emerges as the top-performing category, accounting for a significant share of 56.6% of the total sales/revenue. Following closely is the bag snacks category, responsible for 25.3% of sales.

Delving deeper into the bag snacks category, the chart reveals that pretzels hold a dominant position, contributing almost the entire share of sales within this category. This highlights the immense popularity of pretzels among consumers, making them a key driver of revenue in the bag snacks segment.

The data emphasizes the importance for stores to stock up on high-performing sub-categories to maximize sales potential. Sub-categories like "All Family Cereal," "Pretzels," "Kids Cereal," and

"Frozen Pizza" collectively constitute over 85% of the total units sold. By ensuring sufficient inventory of these popular products, stores can cater to customer demand effectively, maintain customer satisfaction, and boost overall revenue.

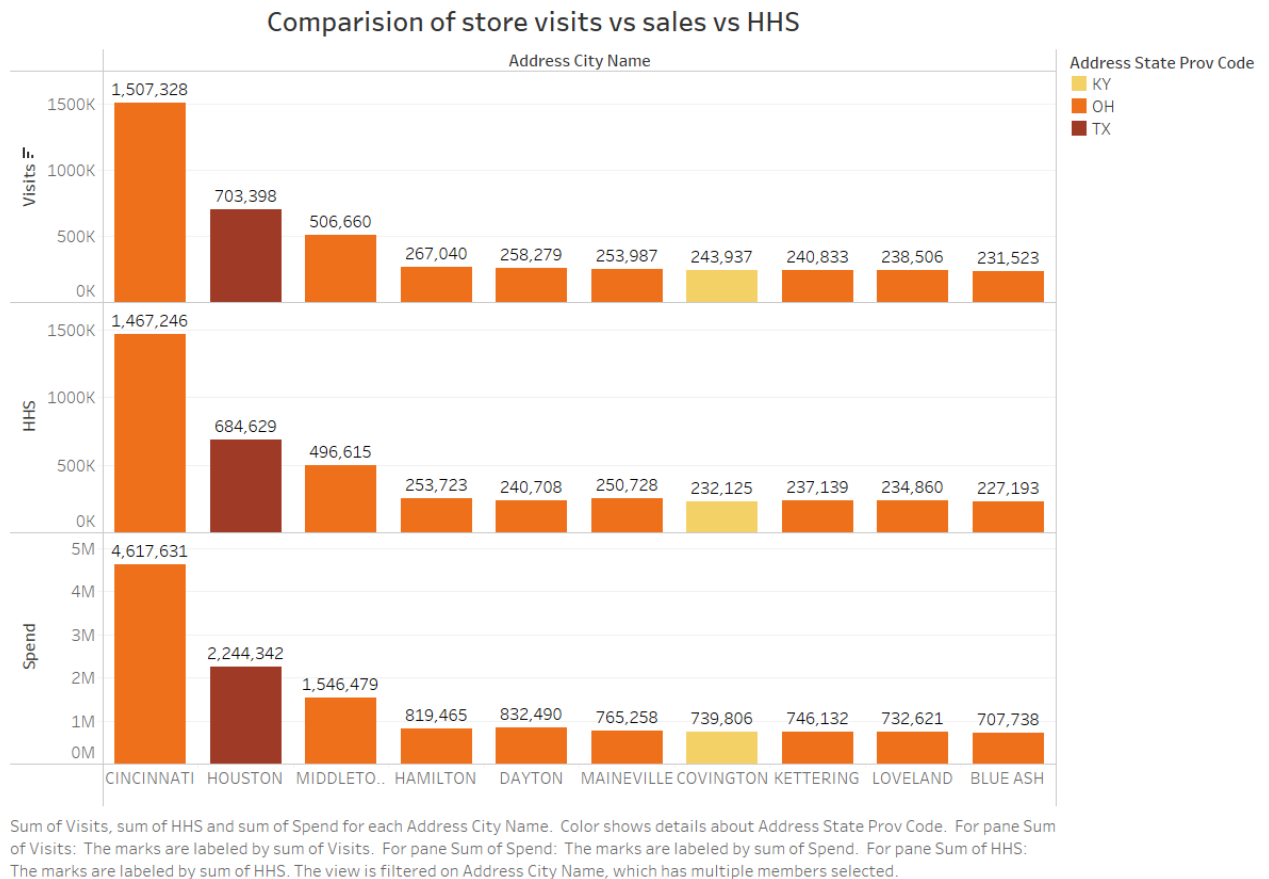
4. Top 10 Cities AVG weekly Baskets



Address City Name and sum of Avg Weekly Baskets. Color shows details about Address State Prov Code. Size shows sum of Avg Weekly Baskets. The marks are labeled by Address City Name and sum of Avg Weekly Baskets. The view is filtered on Address City Name, which has multiple members selected.

The Bubble Graph illustrates the correlation between overall average weekly baskets and sales for all stores in different cities. Cincinnati stands out as the top-performing city with approximately 275k average weekly baskets, closely followed by Houston with around 225k. Notably, the majority of high-performing cities are located in Texas and Ohio, while Covington represents the sole standout city from Kentucky. The graph provides a clear visualization of the cities' performance in terms of average weekly baskets, which directly impacts their sales performance. The data highlights the importance of cities like Cincinnati and Houston in driving revenue, suggesting potential opportunities for businesses to focus on these regions to maximize sales and growth.

## 5. Comparisons of store visits vs sales vs HHS



The above bar chart provides a comprehensive view of the relationship between sales, purchasing households (HHS), and customer visits in various cities and states. By analyzing the data, we can draw valuable insights into the factors driving revenue generation in different locations.

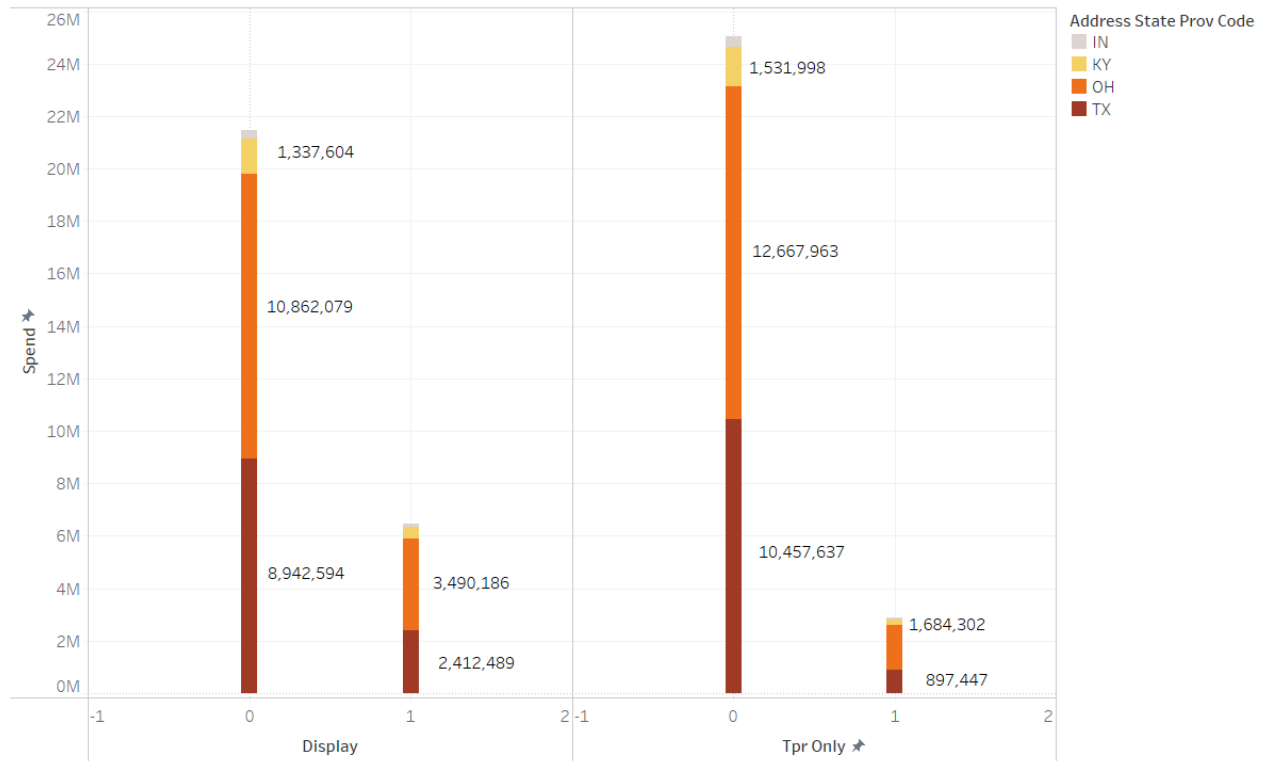
Cincinnati emerges as the top-performing city in terms of revenue, with sales reaching an impressive 4.6 million. This substantial revenue can be attributed to the city's high number of purchasing households and the frequency of customer visits to stores. The correlation between visits, HHS, and sales is evident, indicating that areas with a greater concentration of purchasing households tend to experience higher sales.

Following Cincinnati, Houston, located in Texas, secures the second position with revenue amounting to 2.2 million. The data emphasizes the importance of understanding the local market dynamics, as it plays a significant role in driving sales and revenue.

The insights gained from the bar chart offer strategic guidance for businesses looking to expand their operations or optimize their store locations. Focusing on areas with a high number of purchasing households can lead to a more steady increase in sales. Additionally, understanding customer preferences and behavior in specific regions can help businesses tailor their marketing strategies and product offerings to maximize revenue potential.

## 6. Impact of Promotion & TPR on sales/spends across various regions

Effect of promotions & TPR on sales across various states



The plots of sum of Spend for Display and Tpr Only. Color shows details about Address State Prov Code. The marks are labeled by sum of Spend.

The chart provides insights into the effectiveness of different promotions, such as in-store displays and temporary price reductions (TPR), and their impact on product sales. While some instances showed promising results, in many cases, these promotional methods didn't yield significant improvements. Interestingly, products displayed in-store gained more exposure from customers than those with TPR.

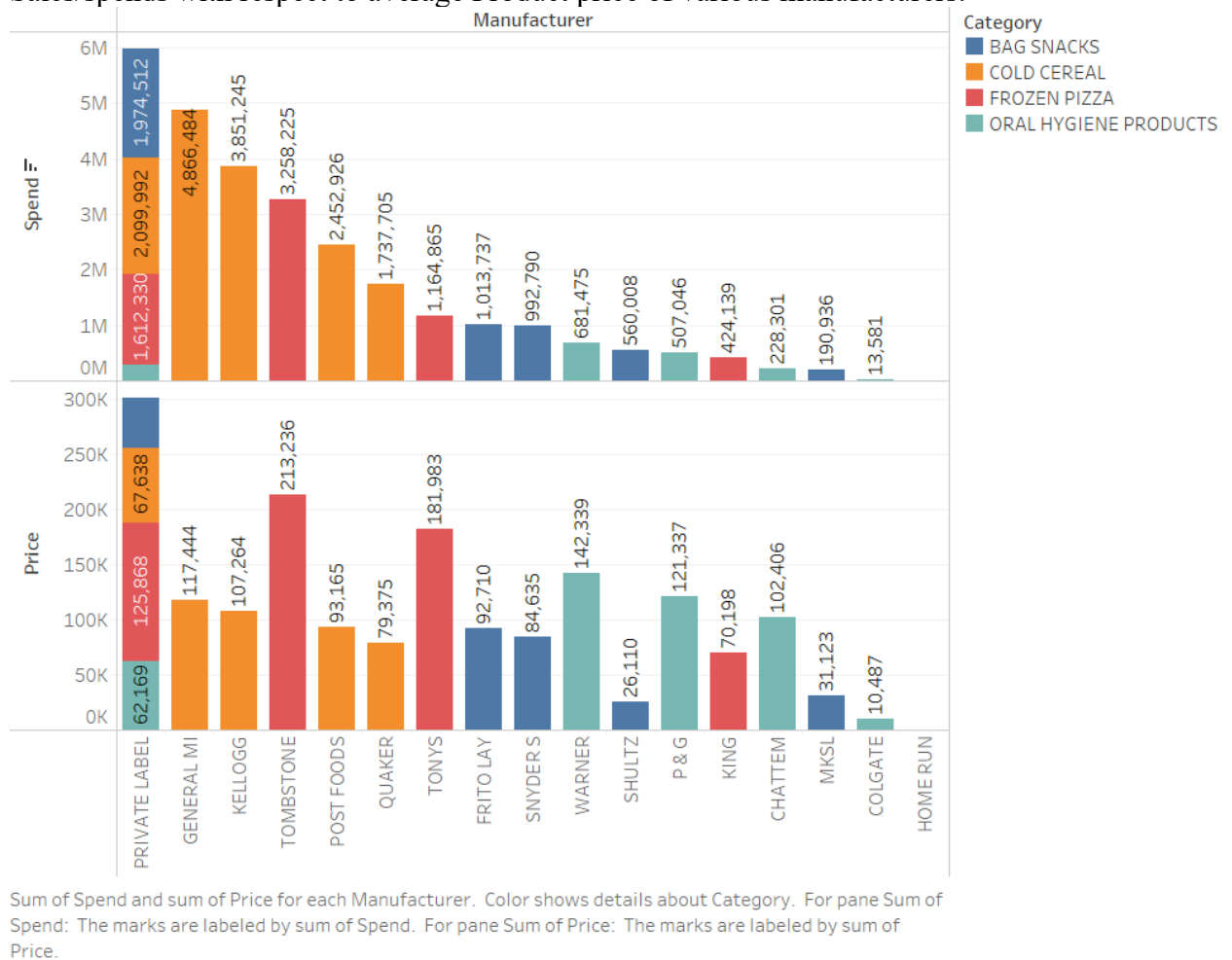
Upon closer examination, it becomes evident that Ohio has shown better results in terms of sales compared to Texas. Specifically, products displayed in Ohio generated sales amounting to 3.49 million, whereas Texas recorded sales of 2.4 million. Even in the case of TPR, Ohio outperformed other states, generating more revenue and achieving a higher ratio of sales with and without TPR.

These findings emphasize the importance of strategically planning promotions based on regional preferences and market dynamics. In-store displays seem to have a more profound impact on customer engagement, while TPR may not always lead to substantial sales improvements. Businesses should consider tailoring their promotional strategies according to regional consumer behavior to optimize sales and revenue generation.

Additionally, these observations can provide valuable insights for retailers and manufacturers to allocate resources efficiently and implement targeted promotions that resonate better with customers in specific states or regions. By leveraging this data-driven approach, businesses can enhance their promotional effectiveness and stay competitive in the market.



## 7. Sales/spends with respect to average Product price of various manufacturers.

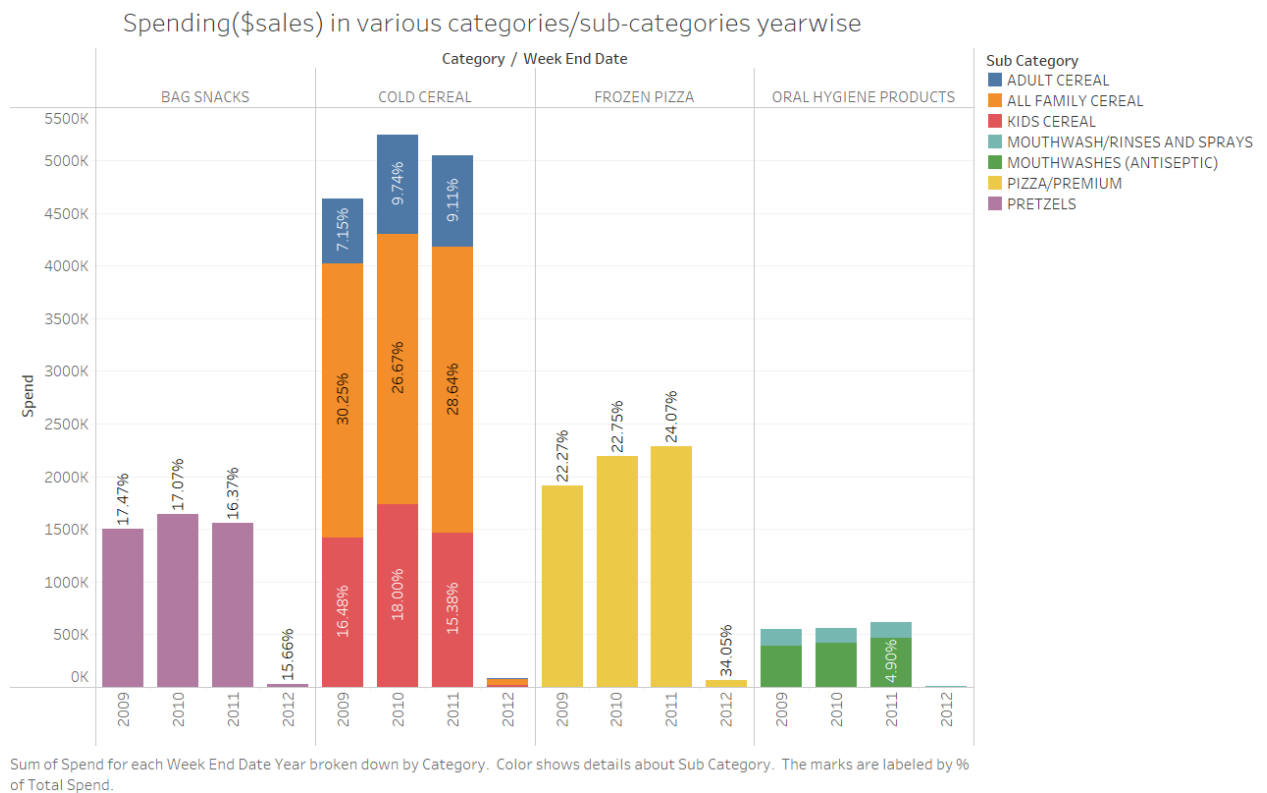


The chart above allows us to observe the trends in the average product price across different manufacturers and their respective sub-categories. Private Label stands out as the top performer with sales of approximately 6 million, spanning products from all four categories. Additionally, by combining the product price with the units sold, we can deduce revenue for each manufacturer.

General MI secures the second position with products from a single category, namely cold cereal, while Kellogg follows closely in third place, also with products in the cold cereal category, generating approximately 3.8 million in sales.

The data provides valuable insights into the performance of various manufacturers and their offerings in specific categories. By understanding these trends, businesses can strategize their product offerings, pricing, and marketing approaches to optimize sales and profitability in the competitive market landscape.

## 8. Spending(sales generated) across various sub-categories annually.



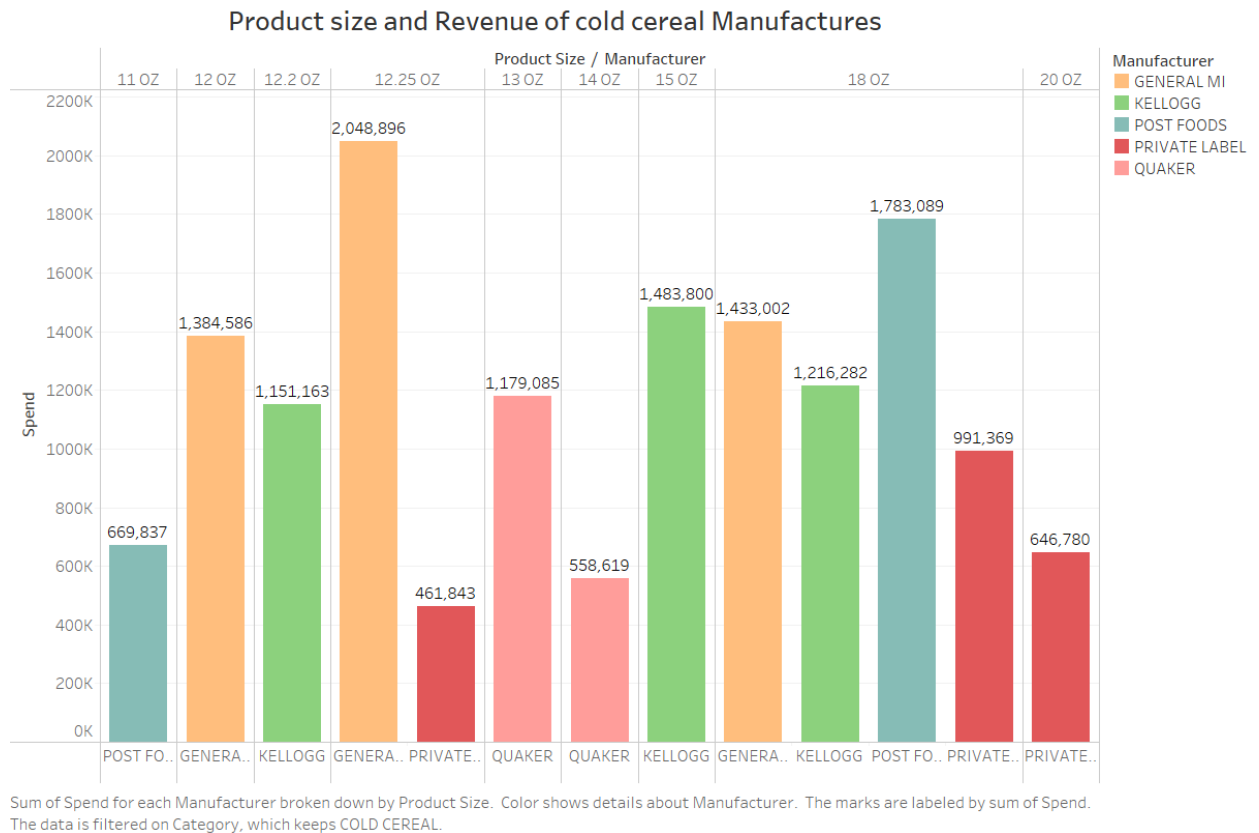
The chart clearly shows the sales trend observed from 2009 to 2012 across various categories and sub-categories offers valuable insights into the market dynamics and consumer preferences during that period. The cold-cereal category consistently dominates in terms of sales, with nearly 5000k in sales each year. However, it is interesting to note that while there was an increase in sales during 2010, the following years experienced a slight decline, indicating fluctuations in consumer demand or market conditions affecting this category.

In contrast, the oral hygiene products category exhibits relatively stable sales over the years, with little to no significant growth. This steadiness suggests that the demand for oral hygiene products remains consistent throughout the given period.

The most remarkable trend, however, is observed in the Frozen Pizza category. This category displays steady growth over the years, making it a standout performer among the various categories. In 2012, the Frozen Pizza category's sales achieved an impressive 34% of the overall category sales, signifying a considerable and sustained demand for frozen pizzas. This growth trend offers a promising opportunity for retailers and manufacturers to capitalize on the increasing popularity of frozen pizzas and explore strategies to further enhance its market presence.

The slight downward trend in most other categories highlights the importance for businesses to stay agile and adapt to changing market conditions. Understanding consumer preferences, responding to market trends, and offering products that align with customer demands are crucial factors for success in the competitive retail landscape.

## 9. Sales of cold cereal manufactures with respect to Product sizes



As we observed in the previous plots, the cold cereal category stands out as a significant contributor to revenue. Now, let's delve deeper into this category by examining the distribution of product sizes among various manufacturers in the bar plot above.

In the 18oz category, all manufacturers offer products, and in the 12.25oz category, only General MI and Private Label compete. Interestingly, all other size segments are monopolized by specific manufacturers.

From this analysis, it becomes evident that General MI leads in selling cold cereals, generating approximately 4.8 million in sales, with Kellogg closely following in second place and Postfoods in third place. Surprisingly, even though Private Label has the highest overall category sales, it falls behind in this specific segment. To improve its sales in the cold cereal category, Private Label should consider strategizing based on the factors observed above.

Understanding the distribution of product sizes and the competitive landscape in the cold cereal category can help manufacturers like Private Label tailor their marketing and product offerings to better compete and capitalize on this lucrative market segment. By leveraging this information, businesses can refine their strategies and position themselves for success in the highly competitive cold cereal market.

All the above worksheets are organized into a story and can be found in the twbx document as it is not possible to visualize a story here. Below is a snip of story page of as a proof.

## Story:

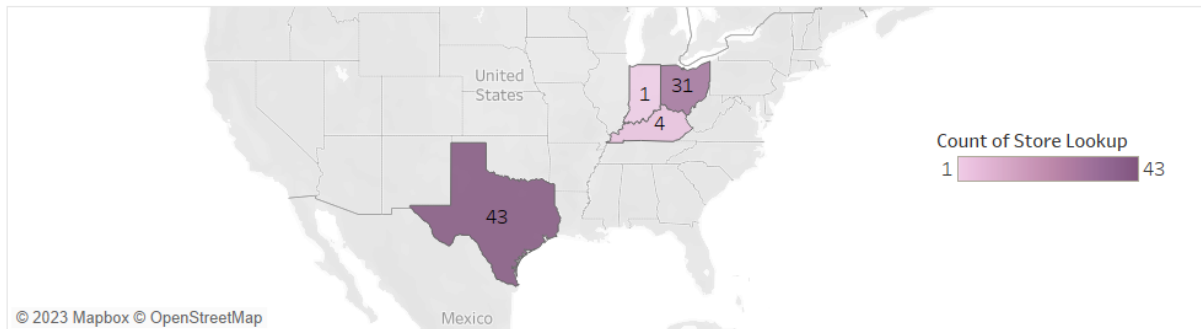
### Story 1

Store locations and manufacturer unit..	Analysis by Manufacturer and ..	Sales & Avg Weekly Baskets analysis	store visits, HHS & Revenue analysis	Impact of overall promotions, TPR's,..	Recommendations
---	---------------------------------	-------------------------------------	--------------------------------------	--	-----------------

#### CollegeLife Data challenge Overview

Count of Products Lookup	Count of Store Lookup	Spend	Units	Visits
58	79	27,927,723	10,293,354	9,012,000

#### Stores in diferent Geographic locations



#### Manufactured units sold in Texas and Ohio



## Recommendations:

1. Private Label should consider extending its product range to include Kids Cereals and Adult Cereals, to expand its company and gain a significant foothold. Additionally, focusing on promoting Mouthwashes, which have a lower average price compared to competitors, can be advantageous. Research indicates that the Kids Cereal Business is profitable, making it a promising addition to Private Label's product lineup, potentially leading to further business growth.
2. Customers prioritize quality over price and quantity, leading to higher sales for high-quality sub-categories. Manufacturers like Kellogg's and Private Label can adjust product pricing based on demand and competition to increase profits. For example, they can reduce prices for less

demanding products and raise prices for popular ones, such as the 18oz size. This strategy can enhance sales and overall revenue.

3. While most other categories have experienced a slight downward trend in sales over the past couple of years, the Frozen Pizza category stands out with consistent growth. Its annual sales have risen from 22.7% to an impressive 34.05%. This indicates a significant and sustained demand for frozen pizzas, making it a standout performer in the market. Retailers and manufacturers should take note of this trend and explore opportunities to capitalize on the increasing popularity of frozen pizzas. Understanding consumer preferences and preferences for convenience and taste may contribute to the category's continuous growth.

4. To increase sales and raise awareness, manufacturers like Shultz and MKSL must focus on effectively advertising their TPRs, especially in the Ohio region where they currently lack sales presence. While they do have a presence in Texas, capturing the Ohio market will require strategic advertising. Utilizing social media and other promotional methods can attract more customers and boost revenue, particularly in highly competitive times. Successful TPR advertising can significantly enhance their market position and help them tap into new opportunities for growth.

5. stores should prioritize stocking up on "All Family Cereal," "Pretzels," "Kids Cereal," and "Frozen Pizza" due to their high popularity, which collectively accounts for over 85% of total units sold. Additionally, it is crucial to understand that customer visits are directly linked to sales. To retain and attract more customers, stores should invest in providing adequate facilities, including potential store size expansion and other improvements. By catering to the demand for these popular sub-categories and enhancing the overall shopping experience, stores can maximize sales and create a more appealing shopping environment for their customers.

**Thank you**