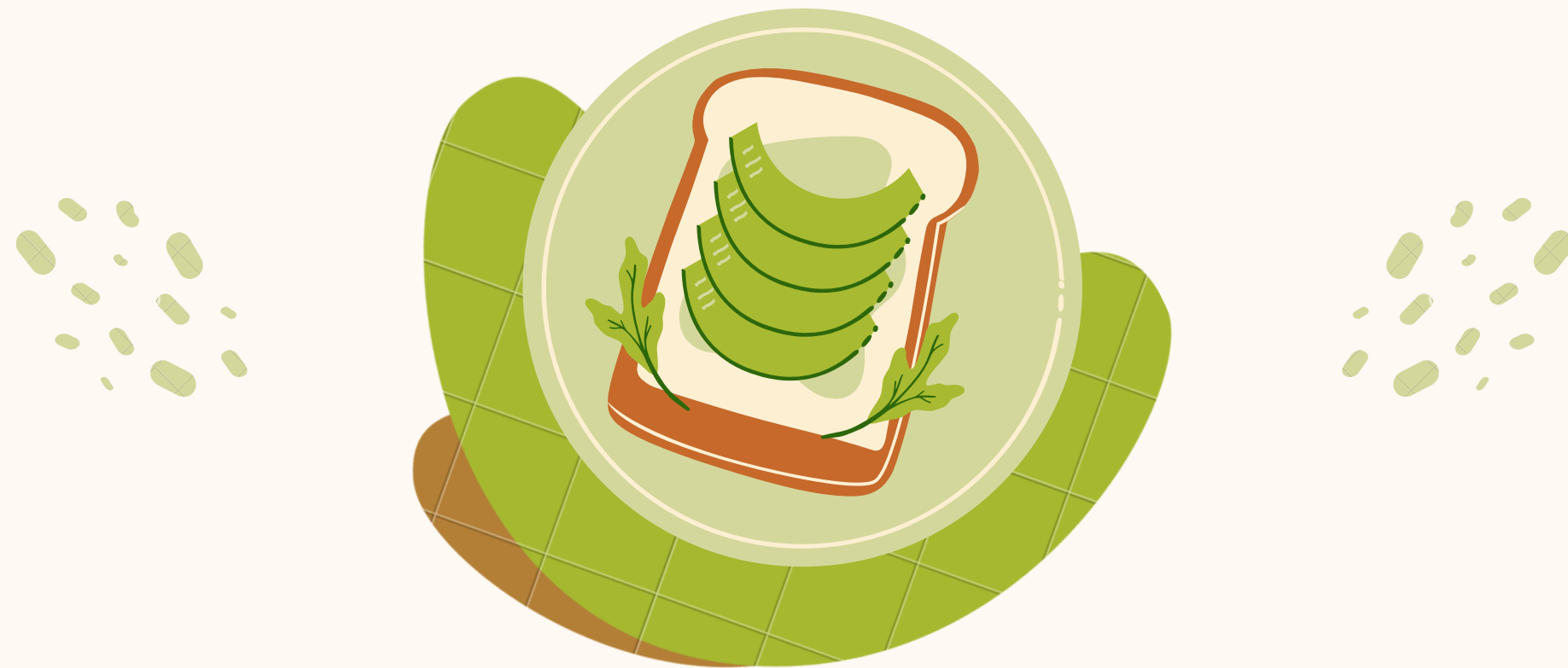# DATA MADE IT BETTER
FOOD RETAILS

Nanoth Tikhinanon

ABOUT DATA

store   store id (744 stores), zip code

causal   feature_desc (how to promote), disply_desc (how to display product in the store)

product   product id, brand, description

transaction   units, product, household (buyer), time_transaction, basket, week, day, store, coupon (discount)

728 days, 104 weeks, 2 years of transactions

GOALS

- More profit
- Fewer unnecessary costs
- Increased product sales
- Improved customer loyalty.

4 SOLUTIONS

Reduce Stock

Ads & display positions

The coupon!

Promotion: buy two, get a better deal!

Nanoth Tikhinanon

REDUCE STOCK

In fact, we have four commodities that offer a variety of brands:

- Pancake mixes: 17 brands
- Pasta: 56 brands
- Pasta sauce: 43 brands
- Syrups: 35 brands

This implies: **more brands, more substitutes**

first

REDUCE STOCK

After analyzing all transactions over two years, several UPCs (SKUs) had minimal units sold. For instance:

- Aunt Jemima - syrups
- Pastariso - pasta
- Tree of Life - syrups

All the mentioned products above had only **one unit sold** over the two-year period.

REDUCE STOCK

Reducing stock for products with unit sales below a certain threshold will **cut unnecessary costs**. Establishing a threshold of 10 units sold could potentially reduce 147 UPCs.

In the future, considering a time-based threshold, products failing to sell over 100 units in a five-year span should be reconsidered for sale.

ADS & DISPLAY POSITIONS

In the causal table, there are two intriguing columns: **feature_desc** and **display_desc**. Here are the values for these valuable columns.

feature_desc	
Interior Page Feature	192,995
Not on Feature	78,955
Front Page Feature	34,049
Wrap Interior Feature	22,203
Back Page Feature	7,397
Interior Page Line Item	6,001
Wrap Back Feature	5,004
Wrap Front Feature	4,768

display_desc	
Not on Display	254,776
Rear End Cap	24,656
In-Shelf	20,617
Secondary Location Display	14,475
Front End Cap	10,178
In-Aisle	9,147
Promo/Seasonal Aisle	5,329
Mid-Aisle End Cap	4,675
Store Rear	4,250
Store Front	2,257
Side-Aisle End Cap	1,012

second

ADS & DISPLAY POSITIONS

Given the absence of a data dictionary, let's assume **feature_desc** pertains to advertising products through media, such as on the front page of a magazine. Meanwhile, **display_desc** refers to how the product is showcased in the store.

ADS & DISPLAY POSITIONS

After clustering into nine groups and concentrating on the cluster with the highest average units sold, we've derived some insights:

- Utilize advertising akin to a front-page feature
- Implement displays resembling the rear end cap

Implementing these strategies will help increase the unit sales of the selected products

THE COUPON!

- - - - - - - - - -

The key to business success lies in **retaining loyal customers**, minimizing churn rates, and implementing strategies to encourage repeat visits to our store! To gain insights into this, let's create some features from raw data:

- Count of baskets from each household (indicating how frequently customers make purchases)
- Average range of days between customer visits
- Total transaction count, coupons used, and units of all purchased products
- Average transaction count, coupons used, and units of all purchased products

	household	basket	time_of_transaction	coupon	units	time_of_transaction_mean	coupon_mean	units_mean	range_day
106757	106758	219	329408	0	301	1504.146119	0.000000	1.374429	3.292237
56868	56869	210	336665	1	410	1603.166667	0.004762	1.952381	3.433333
293394	293395	178	297458	0	544	1671.112360	0.000000	3.056180	4.050562
232512	232513	170	277549	0	264	1632.641176	0.000000	1.552941	4.241176
28954	28955	170	280588	0	337	1650.517647	0.000000	1.982353	4.229412
...	...	...	...	...	...	...	...	...	...
192424	192425	1	1724	0	1	1724.000000	0.000000	1.000000	0.000000
388299	388300	1	844	0	2	844.000000	0.000000	2.000000	0.000000
192417	192418	1	1641	0	3	1641.000000	0.000000	3.000000	0.000000
388301	388302	1	1832	0	1	1832.000000	0.000000	1.000000	0.000000
255013	255014	1	1644	0	1	1644.000000	0.000000	1.000000	0.000000

510027 rows × 9 columns

THE COUPON!

After clustering into 7 groups and honing in on the cluster with the highest and lowest average visits, we've discovered something intriguing:

While all groups exhibit similarities across various features, the significant disparity emerges in the group with the highest average visits, which also shows an average **coupon usage of around 50 times** compared to the group with the least average visits!

more coupon, more visits

third

THE COUPON!

```
group 2: 1449
mean of sum basket: 71.13
mean of sum time_of_transaction: 111128.36
mean of sum coupon: 1.14
mean of sum units: 142.73
mean of sum time_of_transaction_mean: 1577.55
mean of sum coupon_mean: 0.02
mean of sum units_mean: 2.00
mean of sum range_day_mean: 10.24
```

```
group 0: 282235
mean of sum basket: 1.70
mean of sum time_of_transaction: 2573.60
mean of sum coupon: 0.02
mean of sum units: 3.12
mean of sum time_of_transaction_mean: 1541.60
mean of sum coupon_mean: 0.01
mean of sum units_mean: 1.84
mean of sum range_day_mean: 82.80
```

highest average visits

lowest average visits

THE COUPON!

- - - - - - - -

Upon drilling down into the **top three groups with the highest average visits**, an interesting pattern emerged. We found consistent interest in the following products and brands:

- Pasta - Private Label (the specific brand is unidentified, potentially indicating muliple brands within this category and a general customer interest in pasta)
- Pasta sauce - Ragu brand

third

THE COUPON!

- - - - - - - - - - - -

Considering coupons as discounts, is it a good idea to offer **discounts for products customers like**?

Moreover, what if we **displayed** specific brands preferred by customers in locations where our second strategy boosted unit sales?

This approach aims to increase sales and retain customers simultaneously.

PROMOTION

BUY TWO, GET A BETTER DEAL!

Let's shift our focus to understanding the relationship between different types of products. In every transaction, customers often purchase more than one product in a single basket. If we identify which products customers tend to buy together more frequently, we can create promotions to increase overall unit sales

fourth

PROMOTION

- - - - - - - - - - -

BUY TWO, GET A BETTER DEAL!

From association rules, we've discovered some unexpected patterns: **Syrups and pancake mixes are frequently purchased together**, which comes as a surprise. Despite pasta having the highest unit sales in the past 2 years, syrups and pancake mixes have had the lowest unit sales.

This suggests the need for promotions to boost sales of these two products. Offering discounts for purchasing these items together might be an effective strategy
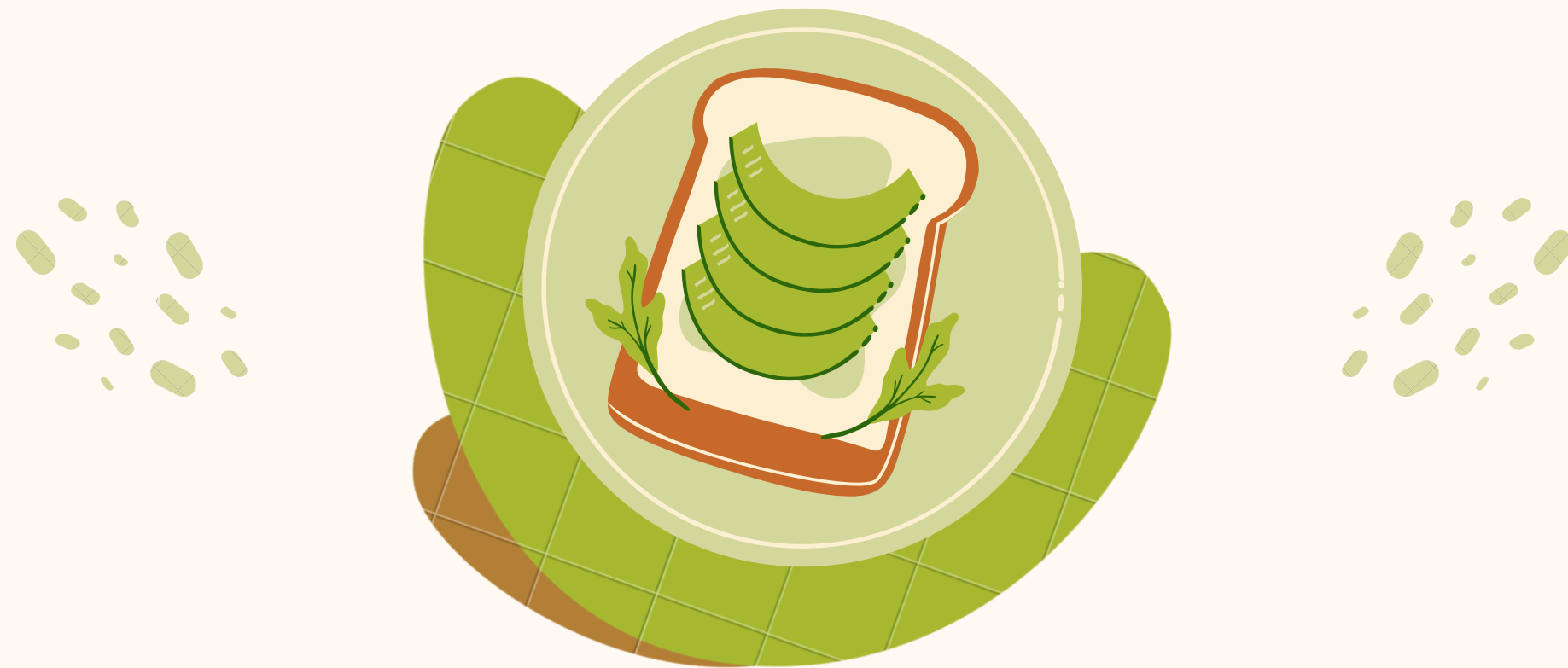
THANK'S FOR WATCH

For further details and Python scripts, you can find them at this link

"Link to the Python scripts"

Nanoth Tikhinanon