MARKET BASKET ANALYSIS CASE STUDY

Data science and machine learning project for analyzing supermarket transaction data with association rule learning techniques.

github.com/leventdusunceli

BUSINESS CASE

A local supermarket wants to learn which items are frequently bought together under different circumstances to optimize their sales approach.

DATASET

The dataset contains 9835 transactions by customers shopping for groceries. The data contains 169 unique items. In our case study we assume that this data belongs to a 3-month period.

OUR APPROACH

In order to identify item combinations, we built Apriori Learning models **for 3 different conditions**. Each of these models will identify the items satisfying the following conditions.

Condition 1

Items that are bought together very very frequently.

Condition 2

Niche items that raise the probability of buying another item drastically

Condition 3

Items that are bought a lot

Condition 1 – Frequently Bought Together

In this condition we wanted to find items that are bought together very frequently and can say that item 1 leads to item 2 with high confidence.

Condition 1 Results

	Left_Hand_Side	Right_Hand_Side	Support	Confidence	Lift
11	(liquor, red/blush wine)	(bottled beer,)	0.001932	0.904762	11.235269
355	(other vegetables, butter, sugar)	(whipped/sour cream,)	0.001017	0.714286	9.964539
51	(misc. beverages, coffee)	(soda,)	0.001017	0.769231	4.411303
186	(bottled water, bottled beer, sausage)	(soda,)	0.001118	0.733333	4.205442
108	(newspapers, spread cheese)	(rolls/buns,)	0.001220	0.750000	4.077529

Comments

75% of customers who bought newspaper and spread cheese also buy rolls/buns.

Chance of soda being bought increases **more than 4 fold** if another alcoholic or non-alcoholic beverage is bought.

Sales Strategy Suggestions

Place spread cheese and newspaper sections in close proximity and put a discount on spread cheese. This will increase the sales of the pair and consequently the sales of buns.

Create a buy 3 get 1 free campaing on bottled beverages and soda. Since people buying beverages are inclined towards buying a soda, to get 1 free they'll buy 2 sodas to get the free beverage.

Condition 2 – Niche Product Combinations

In this condition we identified product combinations that don't appear very often but first part of the combination increases the possibility of second part being bought **more than 5 times.**

For item combos found, effort should be put in **to increase the sales of left hand side** as it will also considerably increase the sale of items in right hand side.

Condition 2 Results

	Left_Hand_Side	Right_Hand_Side	Support	Confidence	Lift
0	(soda, Instant food products)	(hamburger meat,)	0.001220	0.631579	18.995654
4	(soda, popcorn)	(salty snack,)	0.001220	0.631579	16.697793
3	(ham, processed cheese)	(white bread,)	0.001932	0.633333	15.045491
8	(yogurt, hamburger meat, whipped/sour cream)	(butter,)	0.001017	0.625000	11.278670
1	(liquor, red/blush wine)	(bottled beer,)	0.001932	0.904762	11.235269
2	(specialty chocolate, frozen vegetables)	(fruit/vegetable juice,)	0.001017	0.625000	8.645394
17	(other vegetables, herbs, tropical fruit)	(citrus fruit,)	0.001017	0.625000	7.551443

Comments

90% of customers who bought liquor and red/blush wine also bought beer.

Soda popcorn duo makes salty snacks much more attractive to customers, **almost 17 times more** to be exact.

Ham-processed cheese leads to white bread **almost two-third of the time**. Customers seem to like making sandwiches.

Instant food products and soda pair increases the chance of hamburger meat being bought **by almost 19 times**.

Condition 2 – Sales Strategy Suggestions

Create a discount on soda if instant food products are bought. This may increase the sale of hamburger meat.

Put a slight discount on liquor and red/blush wine and a higher price increase to beer. This will increase the sale of liquor-wine combo and since the combo leads to beer purchase 90% of the time, will also lead to increase in profits.

Place ham and processed cheese next to each other and white bread slightly further away. Close proximity will increase the sales of ham-cheese combo.
Once the combo is bought we know there's great chance of a bread purchase so by placing it further we give consumer time to look at other products while walking towards the bread.

	Left_Hand_Side	Right_Hand_Side	Support	Confidence	Lift
0	(soda, Instant food products)	(hamburger meat,)	0.001220	0.631579	18.995654
4	(soda, popcorn)	(salty snack,)	0.001220	0.631579	16.697793
3	(ham, processed cheese)	(white bread,)	0.001932	0.633333	15.045491
8	(yogurt, hamburger meat, whipped/sour cream)	(butter,)	0.001017	0.625000	11.278670
1	(liquor, red/blush wine)	(bottled beer,)	0.001932	0.904762	11.235269
2	(specialty chocolate, frozen vegetables)	(fruit/vegetable juice,)	0.001017	0.625000	8.645394
17	(other vegetables, herbs, tropical fruit)	(citrus fruit,)	0.001017	0.625000	7.551443

Condition 3 – Common Items

In this condition we looked at item combinations that appear very frequently in shopping baskets

For these combinations we can make adjustments knowing that they are top sellers.

Condition 3 Results

	Left_Hand_Side	Right_Hand_Side	Support	Confidence	Lift
0	(root vegetables,)	(other vegetables,)	0.047382	0.434701	2.246605
3	(whipped/sour cream,)	(whole milk,)	0.032232	0.449645	1.759754
1	(root vegetables,)	(whole milk,)	0.048907	0.448694	1.756031
2	(tropical fruit,)	(whole milk,)	0.042298	0.403101	1.577595
4	(yogurt,)	(whole milk,)	0.056024	0.401603	1.571735

Comments

About **5% of all transactions** (more than 4 times a day) contain root and other vegetables together.

Whole milk is combined with sour cream, vegetables, fruits and yogurts very frequently.

Sales Strategy Suggestions

Place whole milk as further away as possible from vegetables, fruits, and yogurt. It seems that 5% of customers buy whole milk regardless of other items. By placing whole milk further away, we'll allow customers to consider buying other products whilst going towards the milk section.

CONCLUSIONS

In this case study we inspected a **supermarket's 3-month transaction data** to find item combinations that are frequently bought together.

We used **association rule learning methods** to analyze the data.

We **found several item combinations** that seem promising under different circumstances and **offered sales strategies** for increasing the sales of these combinations

References

Dataset:
https://www.kaggle.com/datasets/irfanasrullah/groceries/code?resource=download&select=groceries.csv

Project Codes Notebook:
https://github.com/leventdusunceli/Market_Basket_Analysis