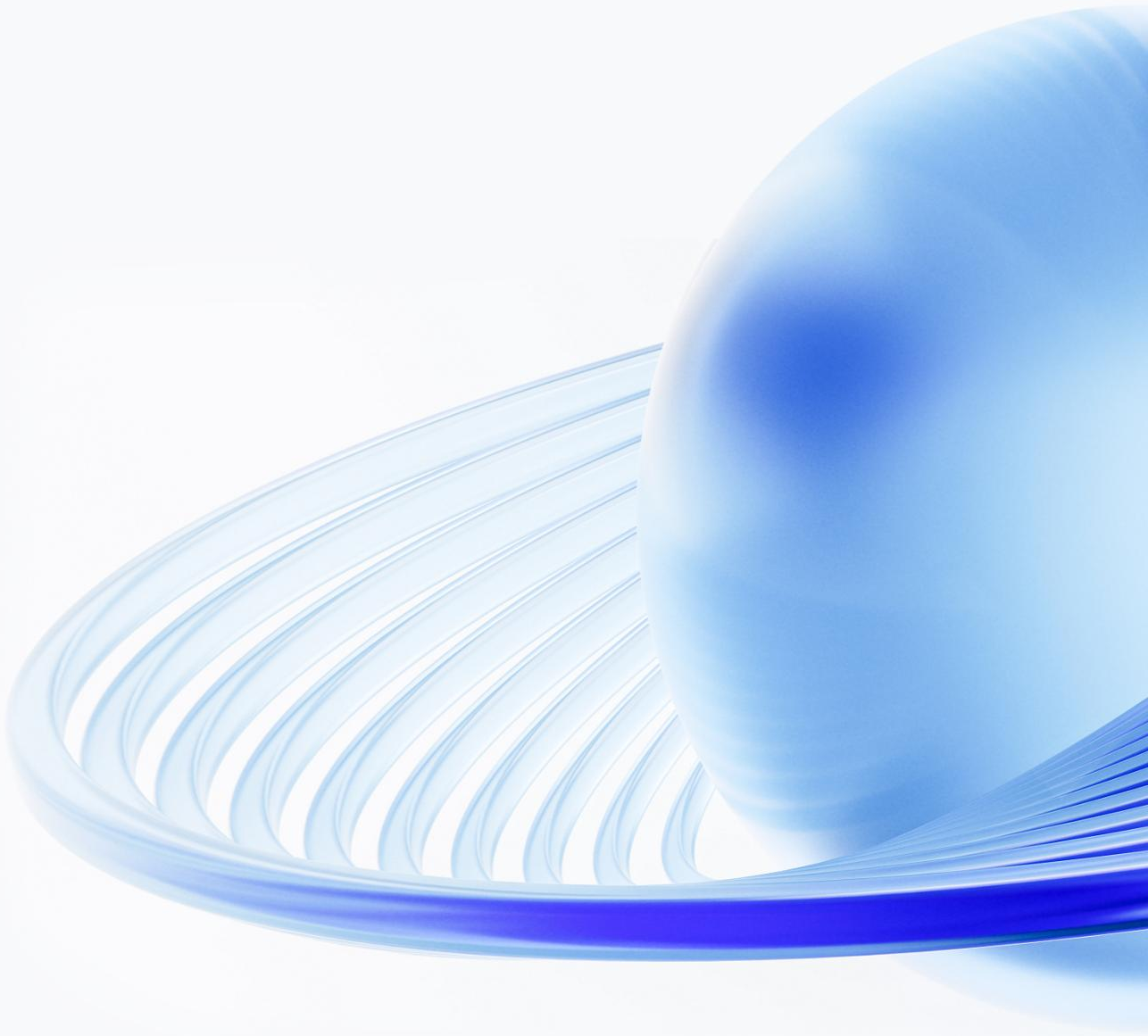
Grocery Store Market Basket

Submitted By:

Akhil Mohandas



Contents

- Exploratory Analysis
- Use of Market Basket Analysis
- Associations Identified
- Suggestion of Possible Combos with Lucrative Offers

Agenda

- Executive Summary of the data
- Exploratory Data Analysis
- Market Basket Analysis
- Associations Identified
- Recommendation

Problem Statement

A grocery store shared the transactional data with you. Your job is to conduct a thorough analysis of Point of Sale (POS) data, identify the most commonly occurring sets of items in the customer orders, and provide recommendations through which a grocery store can increase its revenue by popular combo offers & discounts for customers.

Executive Summary:

- The analysis focuses on transactional Point of Sale (POS) data from 01-01-2018 to 26-02-2020 to uncover patterns in customer purchases. The objective is to provide actionable recommendations for increasing revenue through the implementation of combo offers and discounts.
- Dataset Details
  - Total Records: 20,641 rows
  - Columns: 3 (Date, Order_ID, Product)
  - Missing Values: None
  - Duplicate Values: 4,730
- The exploratory analysis and insights provide a clear understanding of the data and highlight the key trends and patterns in sales.
- **Market Basket Analysis** using association rules was performed to identify the relationships between the products purchased by the customers.
- Dropping duplicate rows may not be appropriate as there is no unique identifier for each row.

Data Dictionary

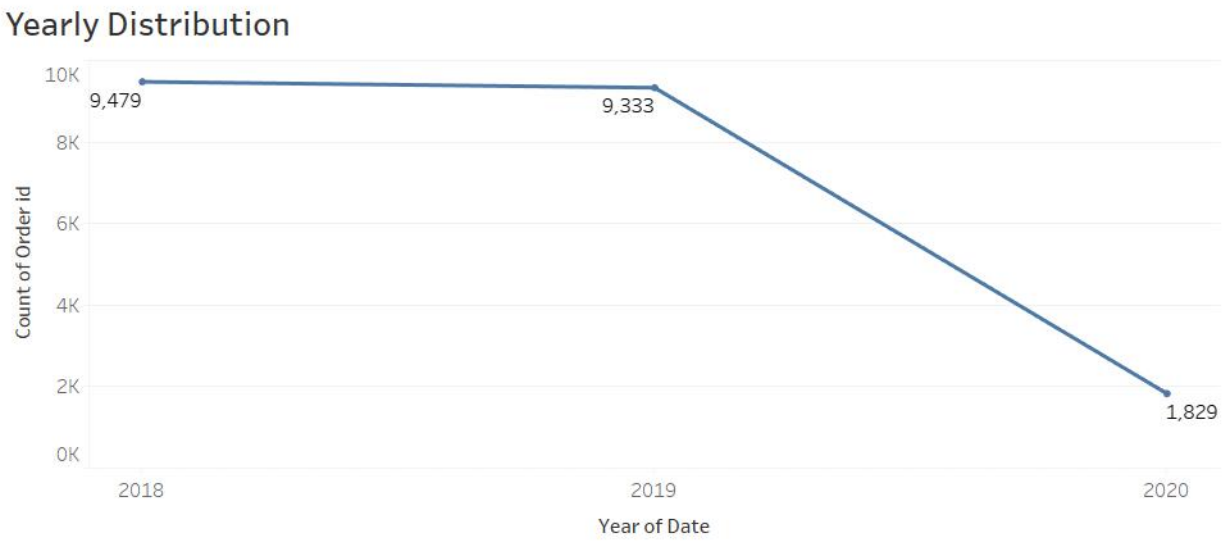
- Date : Date of product sold.
- Order_id : ID of the order.
- Product : Name of the product sold.

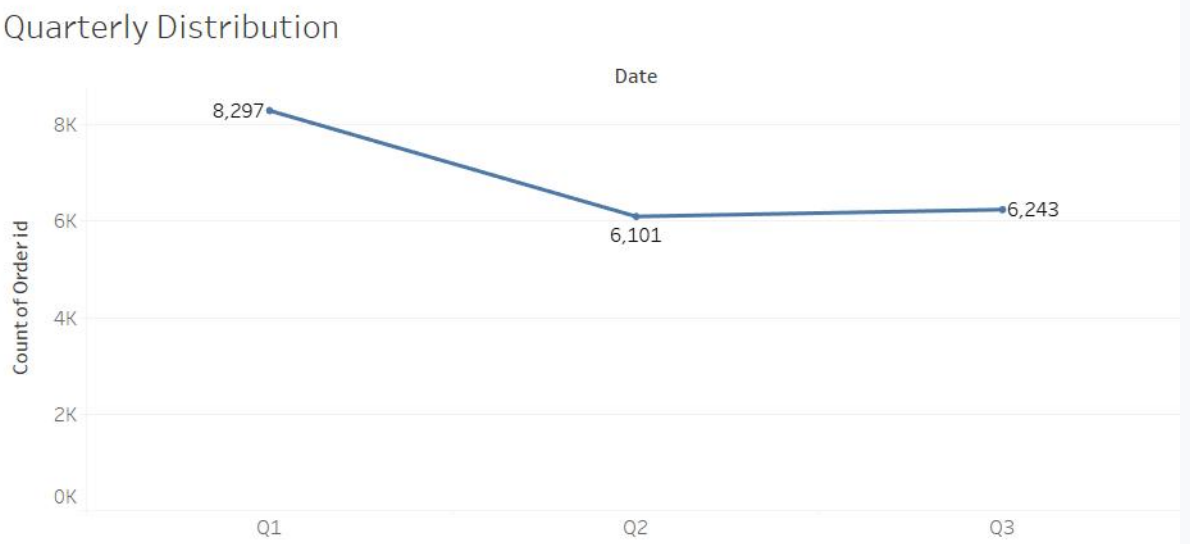
Assumptions

- The dataset represents a list of individual items purchased at a grocery store over a specified period.
- Each row corresponds to a single item purchased during a transaction.
- The first column indicates the date of the purchase.
- The second column represents the Order ID, which uniquely identifies the transaction.
- The third column specifies the item purchased.
- The same item may be purchased by multiple customers across different dates and transactions.
- Transactions are independent, and no information is provided about the sequence or frequency of visits by a specific customer.
- The dataset does not include details such as the quantity or price of each item purchased.
- Analysis focuses solely on identifying patterns in the combinations of items purchased, without financial metrics.
- Duplicate entries were not removed to maintain the integrity of the original data and account for recurring purchases within transactions.

Exploratory Data Analysis

Trend Analysis

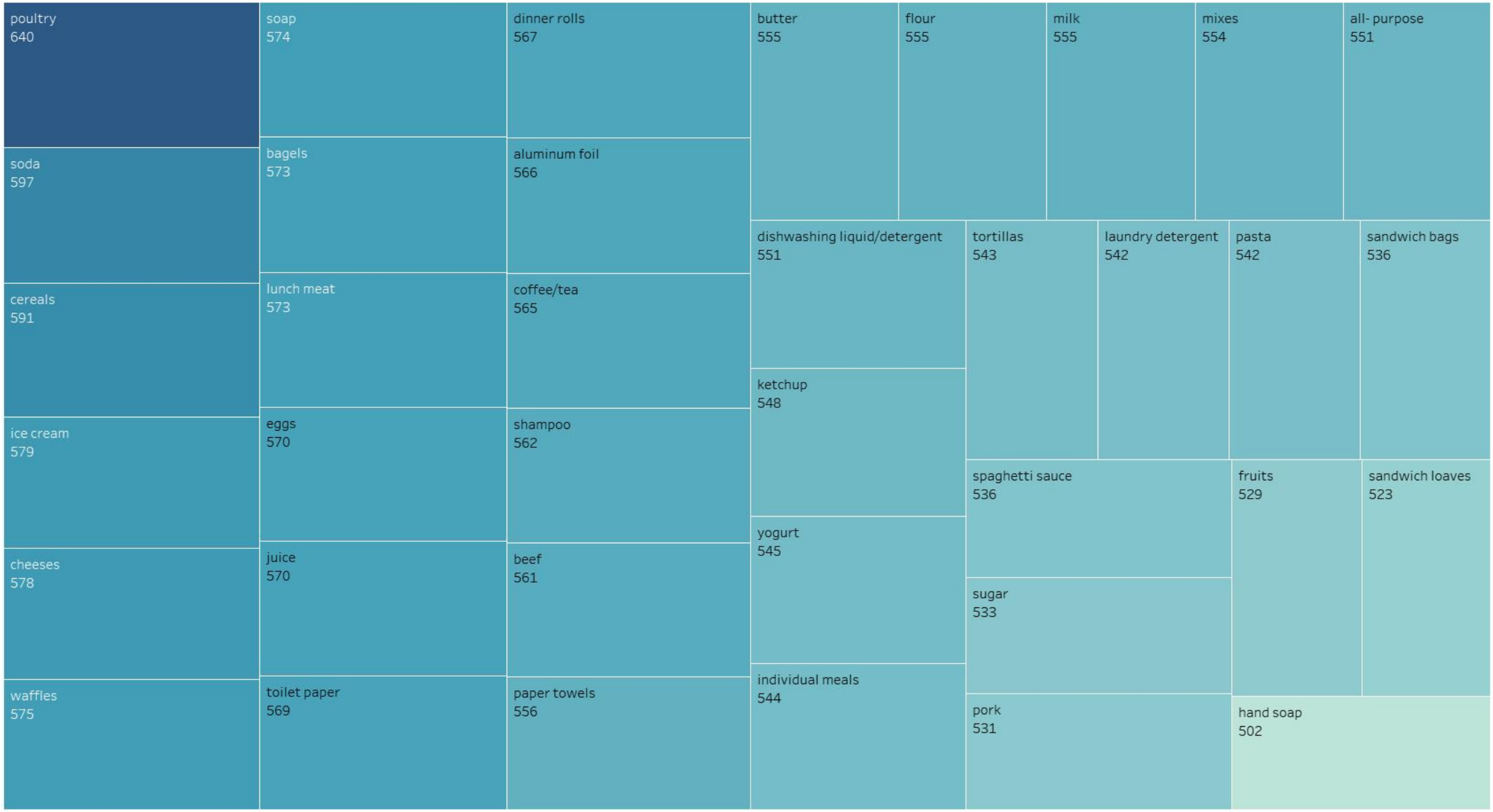




Inferences

- There is a drastic drop in sales for the year 2020.
- Q1 has most sales.
- There is a gradual negative slope in sale from january to april, with a spike in May.
- Lowest products are sold for June month.

Products Sold



Inferences

- Most products sold are poultry, soda and cereals.
- Least sold products are hand soaps, sandwich loaves and fruits.
- Poultry, being the most purchased product, presents an opportunity for strategic promotions, such as discounts or combo offers with complementary items (e.g., spices, vegetables).
- The large range of transactions ensures that the analysis captures diverse customer behaviors and transaction sizes, supporting a robust understanding of purchasing patterns.
- With only 37 unique products, the store can effectively optimize cross-promotions and product placements to maximize sales without overwhelming customers.

Market Basket Analysis

Market Basket Analysis

- Market Basket Analysis is a technique used to uncover relationships or patterns between items purchased together in a dataset. The goal is to identify "association rules," which are if-then statements, such as:

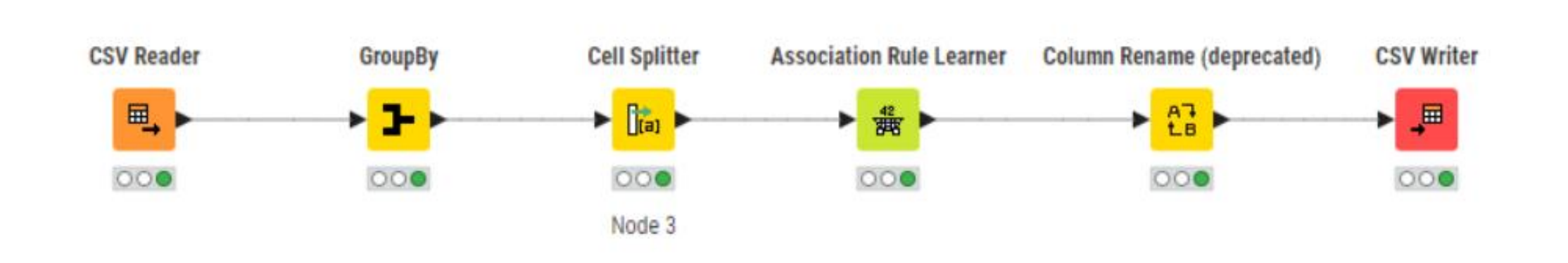
  If a customer buys X, they are likely to buy Y.

- For the grocery store, association rules will help identify frequently bought item sets, enabling the store to:
  - Design combo offers to increase sales.
  - Optimize stock placements to enhance customer convenience.
  - Personalize marketing strategies to target customer preferences.

Metrics

- Support: Indicates how popular the rule is in the dataset.
- Confidence: Shows the likelihood of the consequent item being purchased if the antecedent item is bought.
- Lift: Evaluates the strength of the association beyond random co-occurrence.

MRA KNIME Workflow



Output Table

RowID	Support *Number (double)*	Confidence *Number (double)*	Lift *Number (double)*	Recommended_item *String*	Recommended_with *String*	Items_list *String*
rule0	0.05	0.64	1.7	juice	<---	[yogurt, toilet paper, aluminum foil]
rule1	0.05	0.62	1.645	juice	<---	[yogurt, poultry, aluminum foil]
rule2	0.05	0.613	1.616	coffee/tea	<---	[yogurt, cheeses, cereals]
rule3	0.05	0.6	1.424	poultry	<---	[dishwashing liquid/detergent, laundry deter
rule4	0.051	0.63	1.678	mixes	<---	[yogurt, poultry, aluminum foil]
rule5	0.051	0.611	1.66	sandwich bags	<---	[cheeses, bagels, cereals]
rule6	0.051	0.674	1.726	cheeses	<---	[bagels, cereals, sandwich bags]
rule7	0.051	0.617	1.558	cereals	<---	[cheeses, bagels, sandwich bags]
rule8	0.051	0.63	1.621	dinner rolls	<---	[spaghetti sauce, poultry, cereals]
rule9	0.051	0.637	1.512	poultry	<---	[dinner rolls, spaghetti sauce, cereals]
rule10	0.051	0.604	1.589	milk	<---	[poultry, laundry detergent, cereals]
rule11	0.052	0.628	1.61	eggs	<---	[dinner rolls, poultry, soda]
rule12	0.052	0.641	1.649	dinner rolls	<---	[spaghetti sauce, poultry, ice cream]
rule13	0.052	0.686	1.628	poultry	<---	[dinner rolls, spaghetti sauce, ice cream]
rule14	0.052	0.628	1.614	dinner rolls	<---	[spaghetti sauce, poultry, juice]
rule15	0.052	0.602	1.429	poultry	<---	[dinner rolls, spaghetti sauce, juice]
rule16	0.052	0.634	1.627	eggs	<---	[paper towels, dinner rolls, pasta]
rule17	0.052	0.602	1.621	pasta	<---	[paper towels, eggs, dinner rolls]
rule18	0.054	0.642	1.651	dinner rolls	<---	[spaghetti sauce, poultry, laundry detergent]
rule19	0.054	0.656	1.556	poultry	<---	[dinner rolls, spaghetti sauce, laundry deterç
rule20	0.055	0.624	1.565	ice cream	<---	[paper towels, eggs, pasta]
rule21	0.055	0.63	1.616	eggs	<---	[paper towels, ice cream, pasta]
rule22	0.055	0.643	1.731	pasta	<---	[paper towels, eggs, ice cream]
rule23	0.055	0.649	1.791	paper towels	<---	[eggs, ice cream, pasta]

Recommendation

Inferences

- Poultry and dinner rolls appeared multiple times in association rules, highlighting their strong co-purchase relationships.
- Products like juice, eggs, pasta, and paper towels were identified as complementary items with significant purchase patterns.
- Use the identified combos and discounts to design customer-centric offers.
- Place complementary products near each other in aisles or create combo displays.
- Measure revenue changes and customer uptake on offers to iterate future strategies.

Combos with Lucrative Offers

1. Cheese Lovers' Pack

- Key Products:
  - Recommended Item: Cheeses
  - Associated Items: Bagels, Cereals, Sandwich Bags
- Metrics:
  - Support: 5.09% (indicating the proportion of transactions containing these items).
  - Confidence: 67.4% (likelihood of purchasing cheese when the associated items are bought).
  - Lift: 1.73 (strong association, indicating that this combo is more likely to occur than by random chance).
- Proposed Offer:
  - "Buy bagels and cereals, get sandwich bags at 20% off."
  - Bundle all four items with a 10% discount to encourage bulk purchases.

2. Poultry Feast

- Key Products:
  - Recommended Item: Poultry
  - Associated Items: Dinner Rolls, Spaghetti Sauce, Ice Cream
- Metrics:
  - Support: 5.18% (indicating the proportion of transactions containing these items).
  - Confidence: 68.6% (likelihood of purchasing cheese when the associated items are bought).
  - Lift: 1.63 (strong association, indicating that this combo is more likely to occur than by random chance).
- Proposed Offer:
  - "Buy poultry and dinner rolls, get spaghetti sauce at 30% off."
  - Purchase any three items from this combo, get ice cream free.

3. Household Dinner Combo

- Key Products:
  - Recommended Item: Poultry
  - Associated Items: Dinner Rolls, Spaghetti Sauce, Laundry Detergent
- Metrics:
  - Support: 5.36% (indicating the proportion of transactions containing these items).
  - Confidence: 65.6% (likelihood of purchasing cheese when the associated items are bought).
  - Lift: 1.56 (strong association, indicating that this combo is more likely to occur than by random chance).
- Proposed Offer:
  - "Buy poultry and spaghetti sauce together, get laundry detergent at 25% off."
  - Bundle all four items with a "Buy 3, Get 1 Free" offer.

Thank You