

MARKETING AND RETAIL ANALYSIS

GROCERY STORE

**GAUTHAM GONGADA
POST GRADUATE PROGRAM IN
DATA SCIENCE AND BUSINESS
ANALYTICS**

TABLE OF CONTENTS

01 Basic Data Summary

02 Exploratory Data Analysis

03 Market Basket Analysis

04 Inference from MBA

Problem Statement 2:

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.

OBJECTIVE

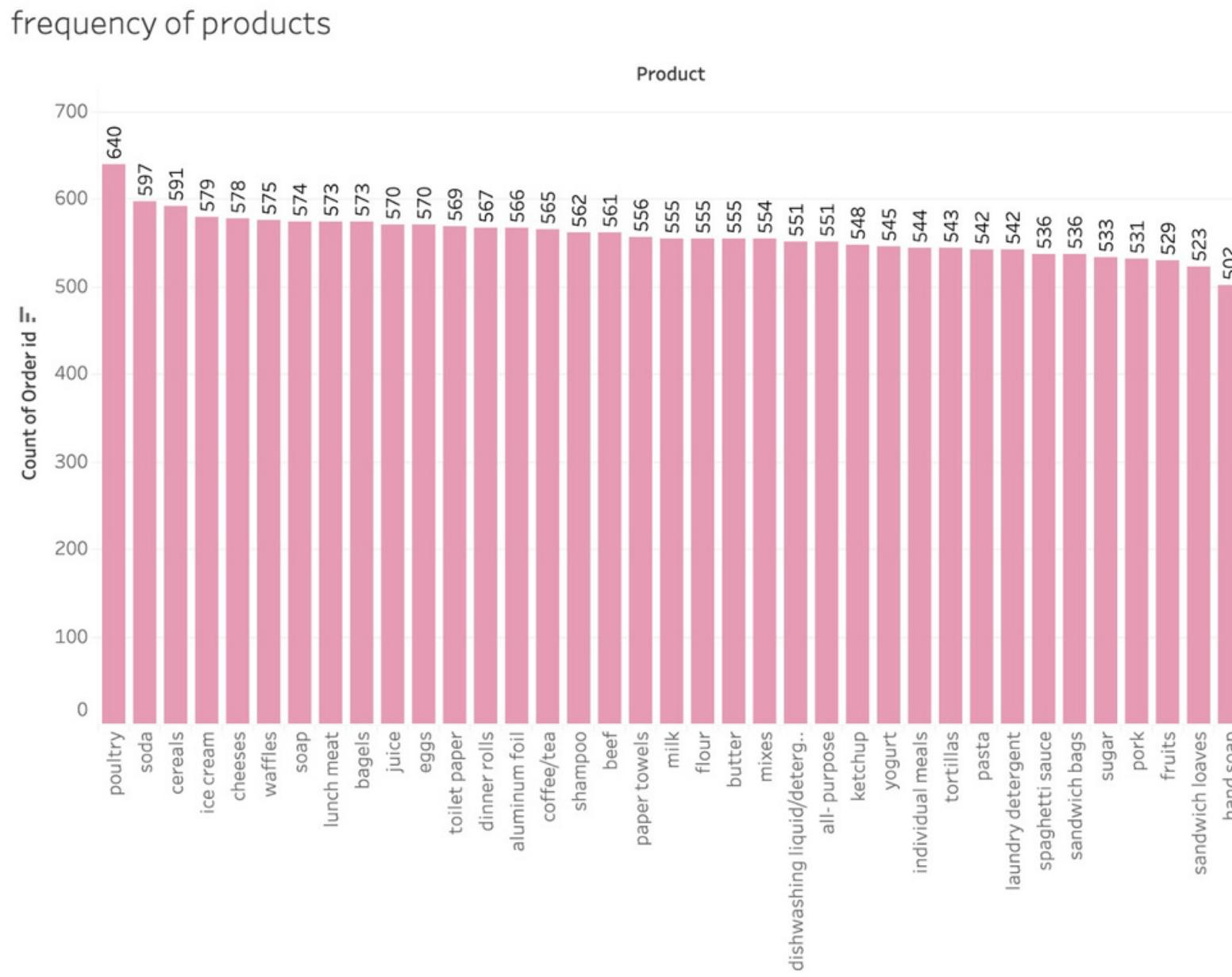
The objective of the project is to conduct a comprehensive analysis of the Point of Sale (POS) data provided by the grocery store. This analysis aims to identify the most frequently occurring sets of items in customer orders. By understanding these common item combinations, the goal is to formulate recommendations for the grocery store to increase its revenue through strategic implementation of popular combo offers and discounts for customers. Essentially, the project aims to leverage data-driven insights to optimize sales strategies, enhance customer satisfaction, and ultimately drive revenue growth for the grocery store.

Basic Data Summary

- The dataset contains records of 20,641 order-ids ordered by customers
- It comprises 3 columns, incorporating 1 integer, and 2 object data types, offering a wide array of information about both transactions and customers.
- There are no null values.
- There are no duplicate entries within the dataset. On average, customers ordered 18 items per transaction, with a minimum of 3 and a maximum of 34 .
- There are in total of 1139 order-ids

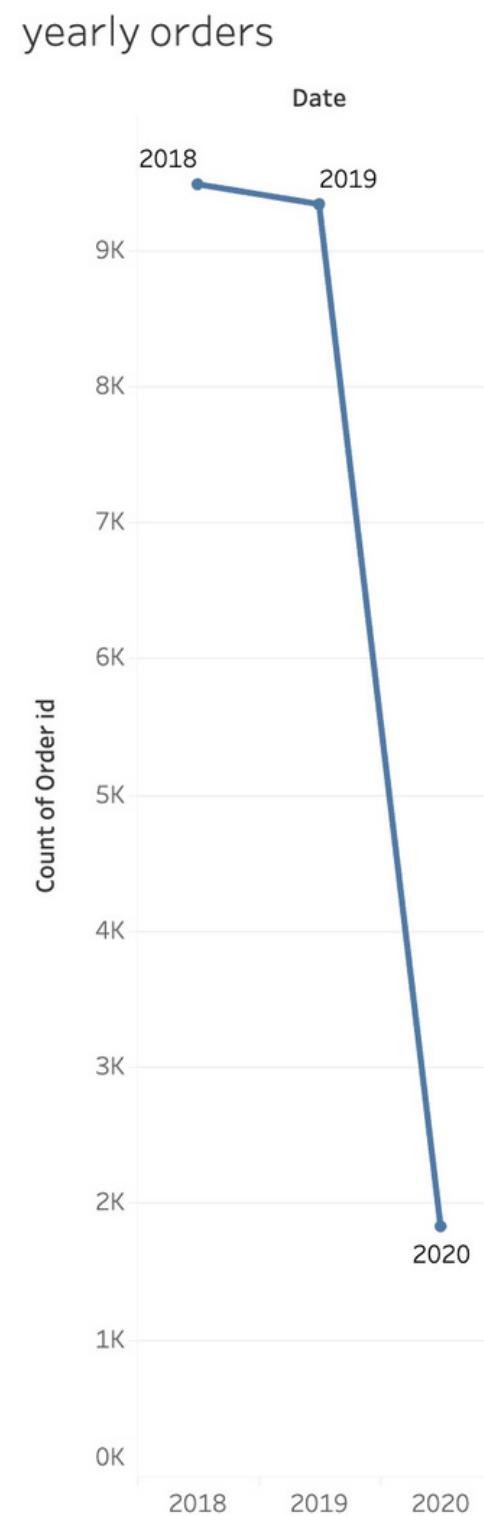
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20641 entries, 0 to 20640
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Date        20641 non-null   object 
 1   Order_id    20641 non-null   int64  
 2   Product     20641 non-null   object 
dtypes: int64(1), object(2)
memory usage: 483.9+ KB
```

Frequency of Products



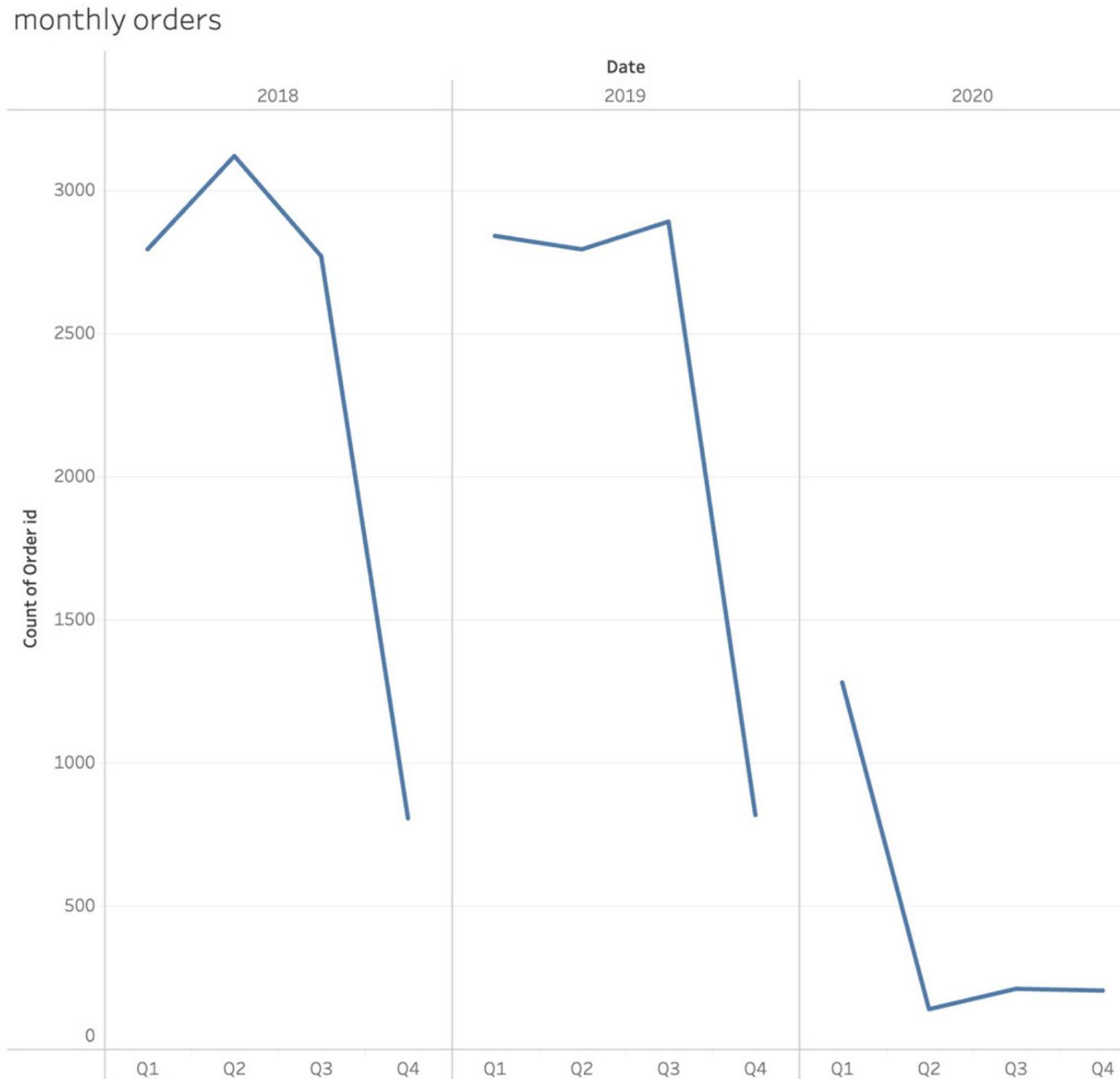
Most purchased products are from poultry category contributing to 31% and least contribution is from hand soaps with 24%

Yearly Orders



There were close to 9.5k orders made in 2018 ,slightly decreasing to 9.2k orders in 2019 and with a sharp decline to 2k orders in 2020

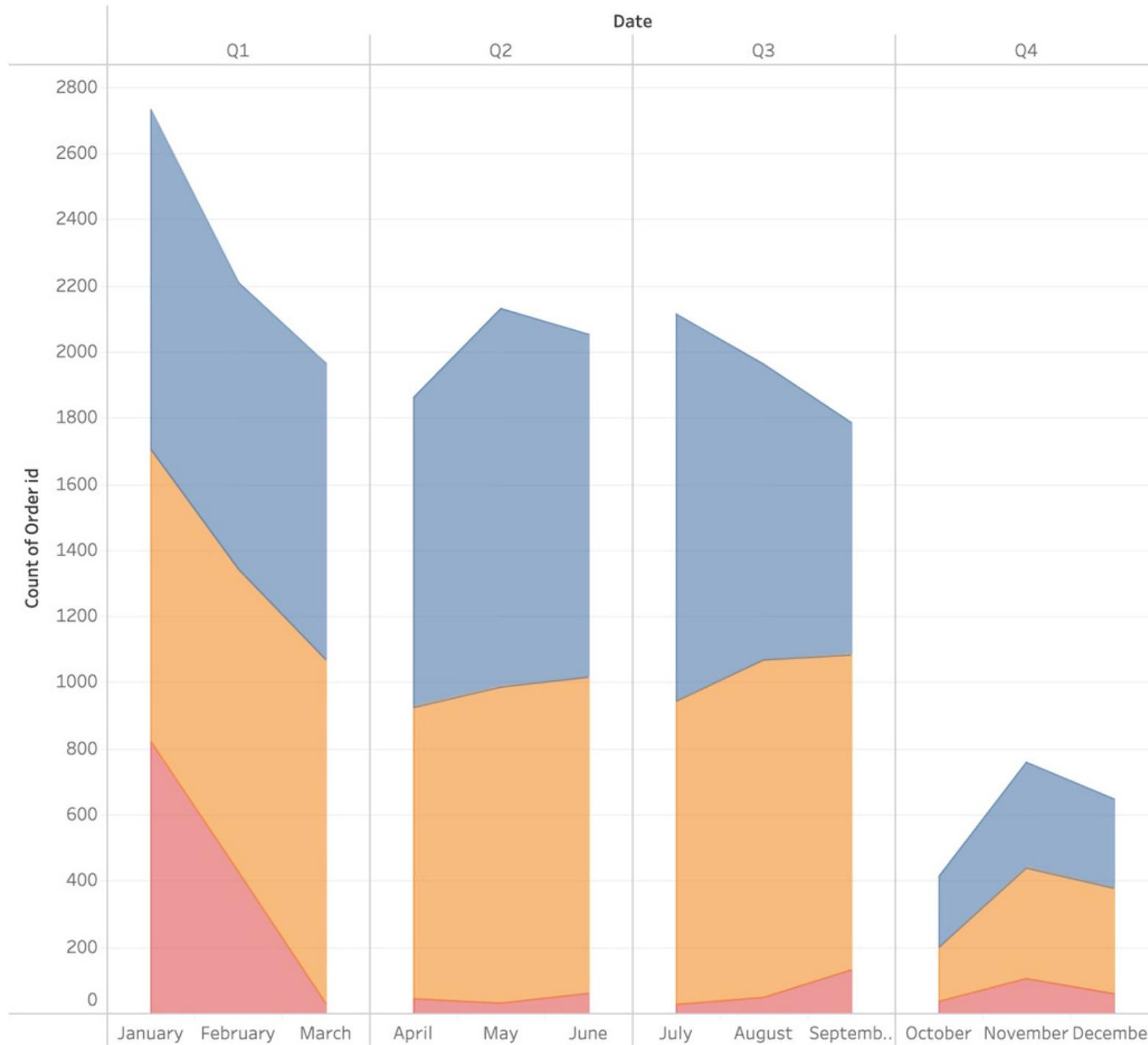
Quaterly Orders



From the data observed, Quarter 1 is having the highest orders .Quarter 2 and 3 are almost having the same amount of orders .Quarter 4 is the least time to receive orders

Monthly Orders

quarterly orders



From the data observed, January is having the highest orders . with June and July having almost the same amount of orders . Months Nov-Dec having the least orders of all the months

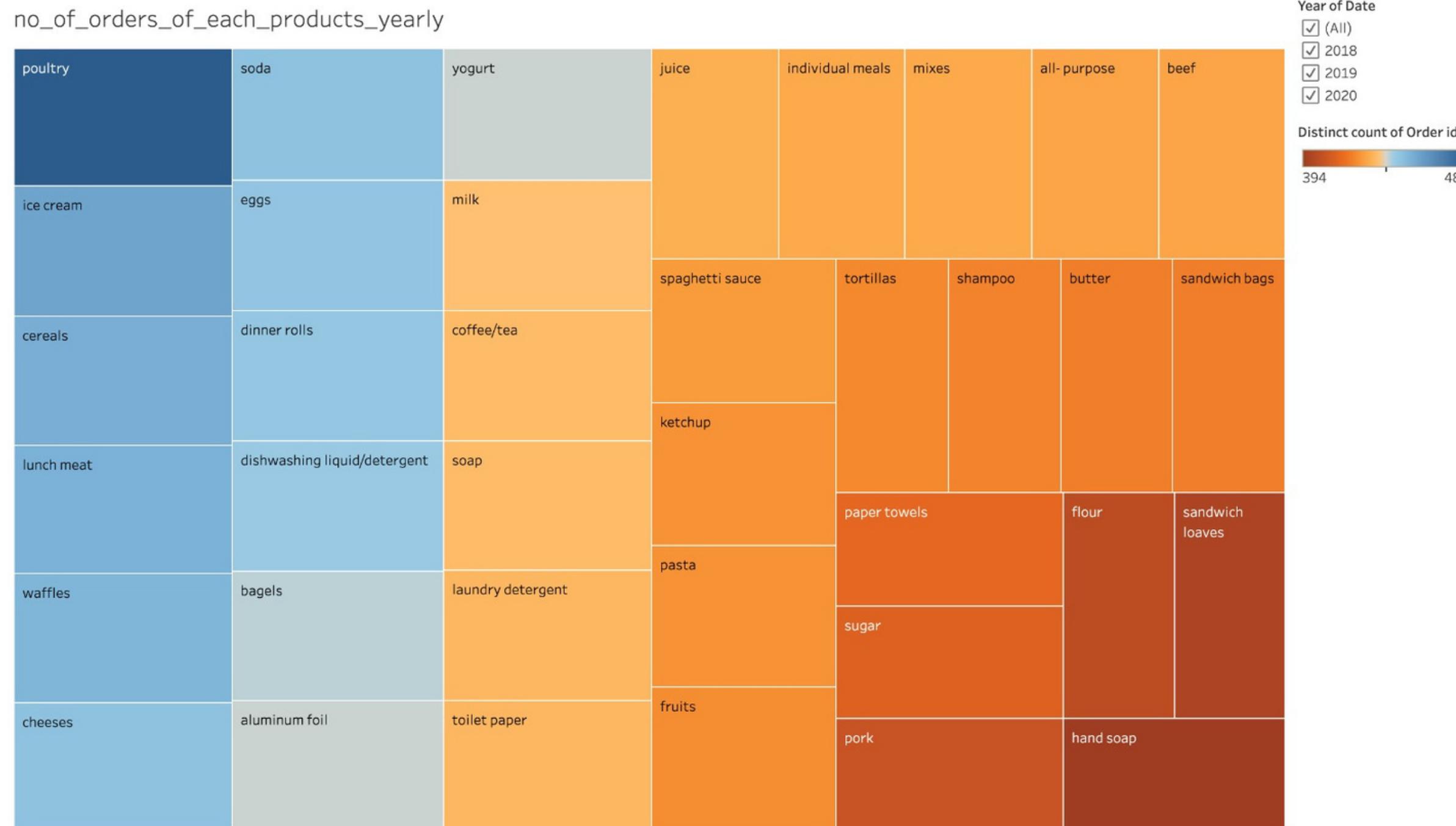
Daily Orders



From the data observed, Quarter 1 was seen more orders in the second half of the month .Quarter 2 has seen the peak in the mid of the month .Quarter 3 has seen the peak at the start of the month.Quality 4 has the highest sales at the end of the month

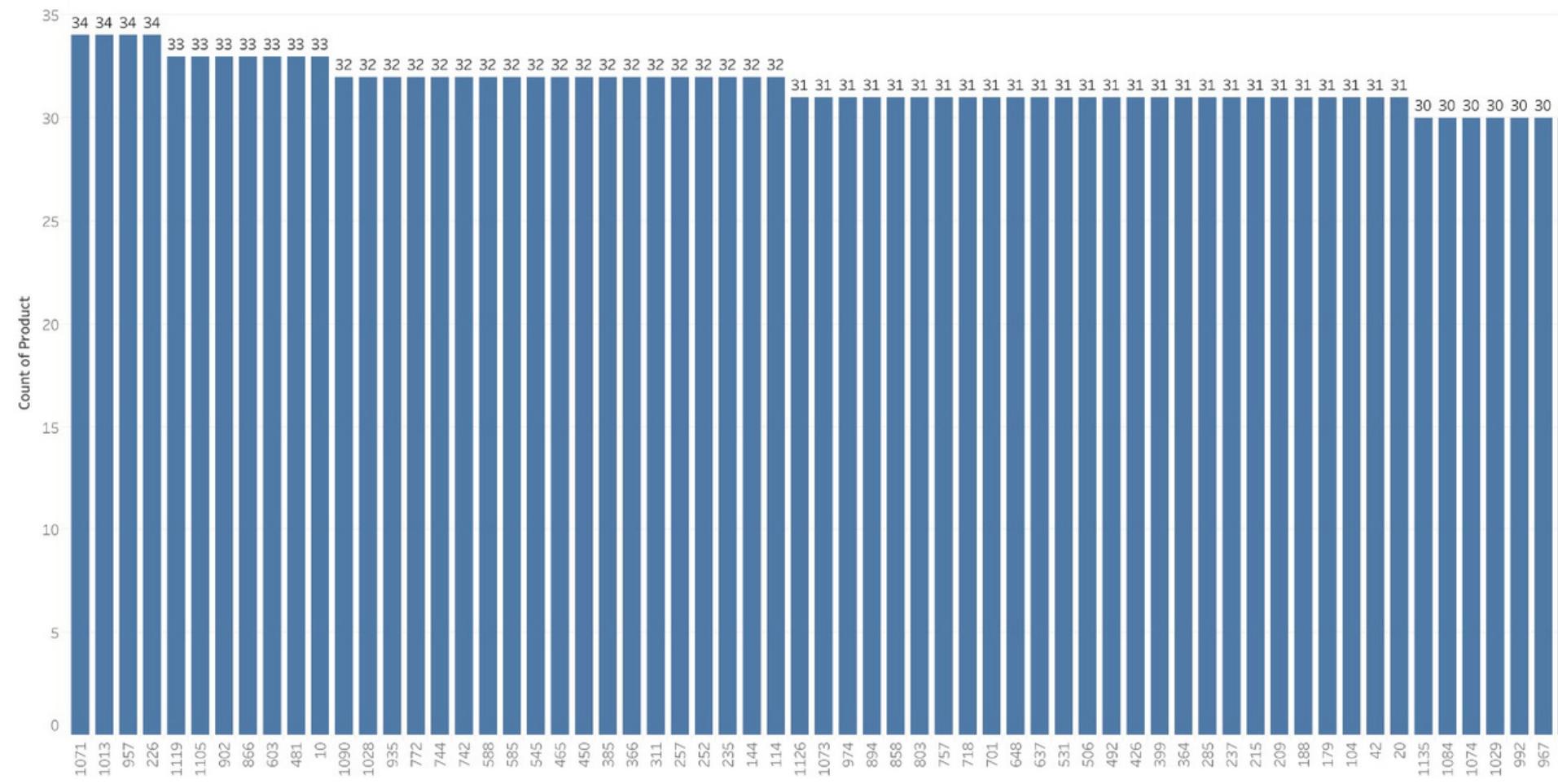
Product count Yearly_

no_of_orders_of_each_products_yearly



Count of product vs order id

count_of_products vs order_id



The highest number of products in a order-id is 34 from 4 different order ids ,whereas the least number of products are 3 from 2 different order-ids

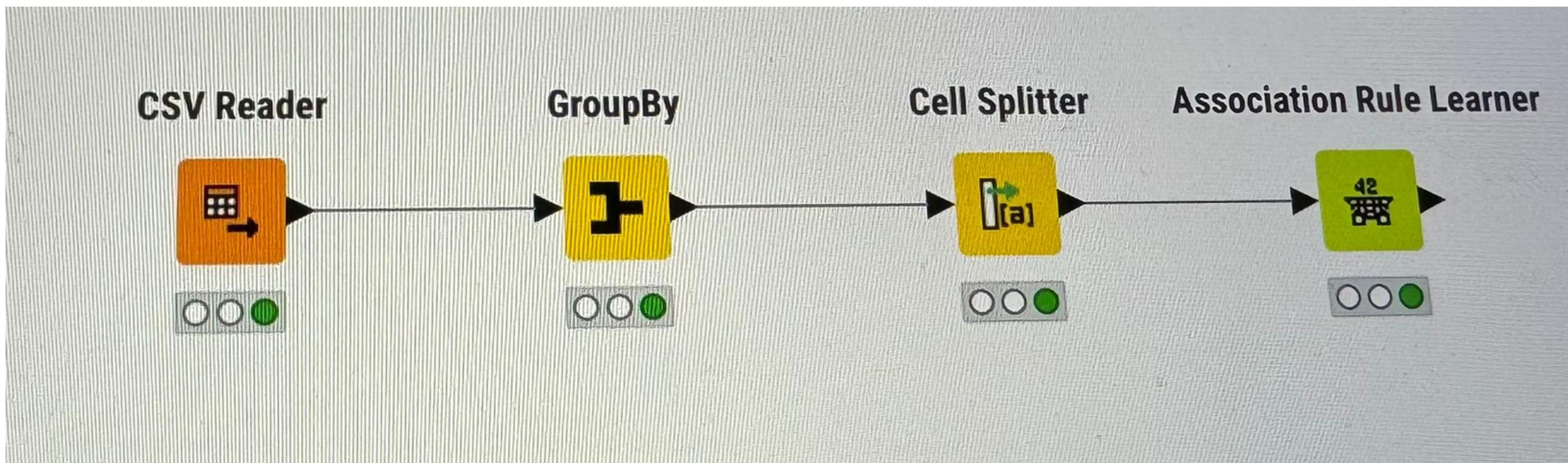
Market Basket Analysis

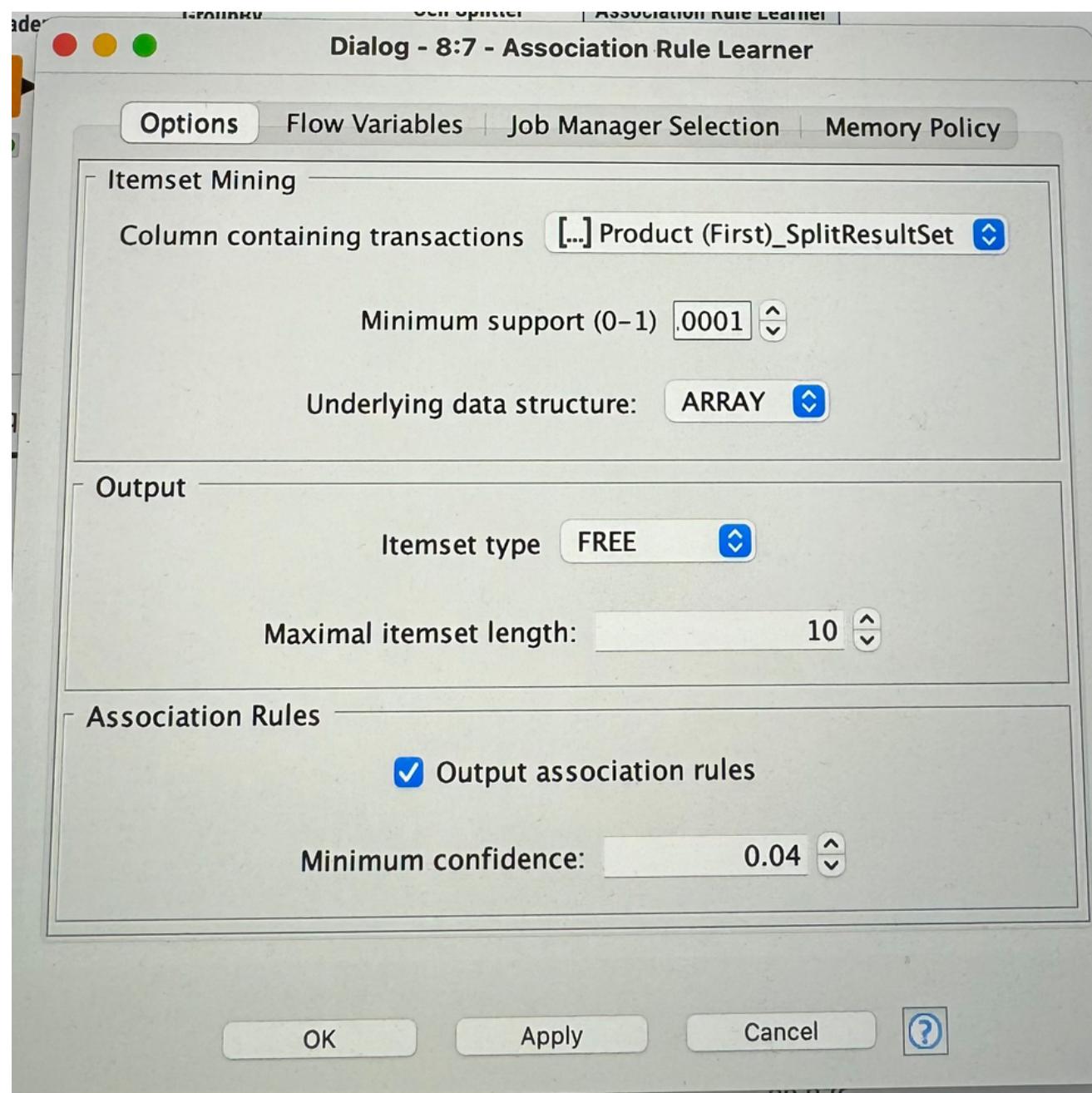
Market Basket Analysis is a method utilized to uncover patterns and connections among items frequently bought together in customer transactions. It assists in comprehending customer behavior and unveiling product relationships, aiding businesses in refining sales strategies and boosting revenue. Through the examination of transactional data, Market Basket Analysis unveils insights on item co-occurrence and association rules, empowering businesses to make informed choices regarding product bundling, cross-selling, and targeted marketing initiatives. This analysis offers valuable understandings into customer preferences, enabling businesses to enhance customer satisfaction and propel business expansion.

Association rules

- In Market Basket Analysis, association rules unveil the connections and occurrences among items, offering valuable insights into customer buying patterns and preferences. These rules are instrumental in assisting businesses with improving product positioning, devising focused marketing initiatives, and executing successful cross-selling and upselling tactics to boost customer contentment and drive revenue growth

Knime Workflow





The support threshold is set to as lowest as 0.0001 to get the desired no. of records

Rows: 28 Columns: 6							
#	RowID	Support	Confidence	Lift	Consequence	Implies	Items
		Number (double)	Number (double)	Number (double)	String	String	Set
1	rule0	0.021	1	47.458	individual	<--	[meals]
2	rule1	0.021	1	47.458	meals	<--	[individual]
3	rule2	0.023	1	24.234	soap	<--	[hand]
4	rule3	0.023	0.553	24.234	hand	<--	[soap]
5	rule4	0.025	1	18.672	paper	<--	[towels]
6	rule5	0.025	0.459	18.672	towels	<--	[paper]
7	rule6	0.025	1	39.276	laundry	<--	[detergent]
8	rule7	0.025	1	39.276	detergent	<--	[laundry]
9	rule8	0.025	1	39.276	dinner	<--	[rolls]
10	rule9	0.025	1	39.276	rolls	<--	[dinner]
11	rule10	0.025	1	39.276	all-	<--	[purpose]
12	rule11	0.025	1	39.276	purpose	<--	[all-]
13	rule12	0.026	1	37.967	spaghetti	<--	[sauce]
14	rule13	0.026	1	37.967	sauce	<--	[spaghetti]
15	rule14	0.029	0.541	18.672	toilet	<--	[paper]
16	rule15	0.029	1	18.672	paper	<--	[toilet]
17	rule16	0.029	1	34.515	lunch	<--	[meat]
18	rule17	0.029	1	34.515	meat	<--	[lunch]
19	rule18	0.029	1	34.515	dishwashir	<--	[liquid/detergent]

Support, Confidence, and Lift

- Support, Confidence, and Lift are key metrics used in Market Basket Analysis to understand item relationships and customer purchasing behavior.
- Support indicates how frequently a combination of items appears together in transactions. It shows the popularity of an itemset in the dataset.
- Confidence measures the probability that if a customer buys one item, they will also purchase another item. It assesses the likelihood of co-occurrence between items in transactions.
- Lift determines the strength of association between two items in an association rule. It compares the probability of the items being purchased together to the probability of them being bought independently. A lift value above 1 signifies a positive association, indicating that the items are more likely to be purchased together.

- Confidence measures the reliability of the association rule between two itemsets.
- It is calculated as the proportion of transactions containing both the antecedent and consequent of the rule, out of the transactions containing the antecedent.
- Higher confidence values indicate stronger associations between itemsets.
- By setting a confidence threshold, users can focus on rules that are highly likely to occur.
- Like the support threshold, the choice of confidence threshold depends on the dataset and the desired level of certainty in the association rules.
- In KNIME, users typically set a minimum confidence threshold as a percentage.
- Choosing the right values for these thresholds involves a trade-off between the comprehensiveness of the analysis and the specificity of the results. If the thresholds are set too low, it may result in a large number of irrelevant rules, while setting them too high may lead to missing out on potentially interesting associations. Therefore, it's often an iterative process, where analysts adjust the thresholds based on the insights gained and the business requirements.

Recommendations

- Here are some simple recommendations to boost sales based on item associations:
- Combo Deal: Offer customers a special combo deal where they can buy yogurt, poultry, and aluminum foil along with juice at a discounted price or with an additional item.
- Buy Two Get One Free: Introduce a "buy two get one free" offer on dinner rolls, spaghetti sauce, and ice cream to encourage customers to purchase these items together.

These recommendations leverage item associations to enhance customer satisfaction and encourage exploration of additional products.

- **Bundle Promotion:** Create a bundle promotion where customers can purchase paper towels, eggs, and pasta together at a discounted rate.
- **Cross-Selling Offer:** Provide a cross-selling offer where customers buying cereals can receive a discount on cheese, bagels, and sandwich bags.
- **Limited-Time Promotion:** Launch a limited-time promotion where customers purchasing poultry, laundry detergent, and mixes can enjoy a percentage discount.

Loyalty Program: Implement a loyalty program rewarding customers who frequently buy recommended items or participate in suggested combos with exclusive discounts or rewards. This encourages customer loyalty and repeat purchases.

These recommendations leverage item associations to enhance customer satisfaction and encourage exploration of additional products.

