Marketing and Retail Analysis- Problem 2:

Grocery store data

-Apoorva p

Contents

1. **Agenda & Executive Summary of the data**
2. **Exploratory Analysis and Inferences**
3. **Market Basket Analysis**
4. **Associations Identification**
5. **Suggestion of Possible Combos with Lucrative Offers**

Agenda & Executive Summary of the data

Problem Statement

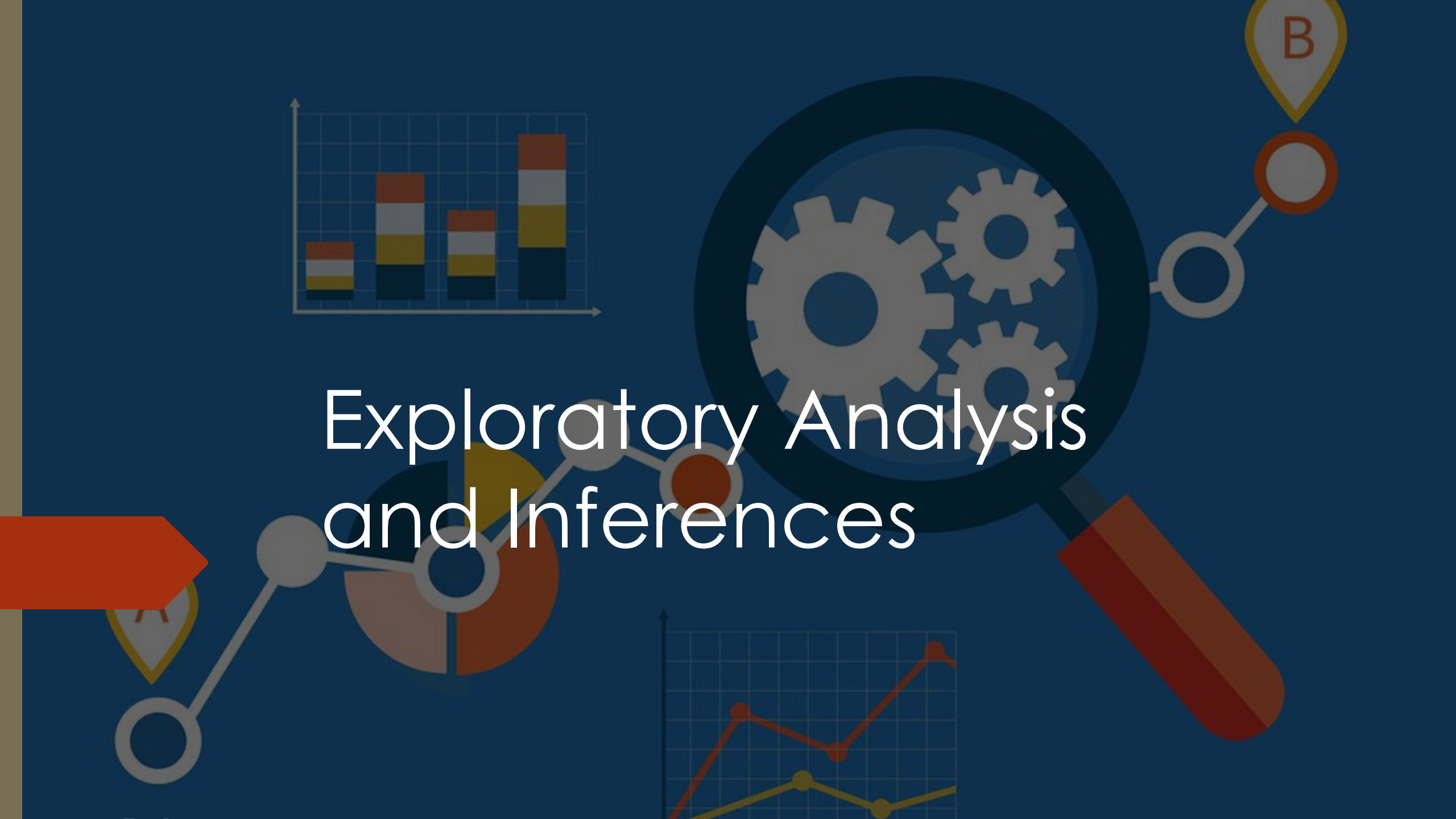
- A grocery store shared the transactional data with you. Your job is to conduct a thorough analysis of Point of Sale (POS) data, identify the most commonly occurring sets of items in the customer orders, and provide recommendations through which a grocery store can increase its revenue by popular combo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

About the data

- The data consist of 20641 rows and 3 columns.
- The data doesn't have any null values, although it has duplicate values in order id but as they signify buying of different product it will be considered

OBJECTIVE

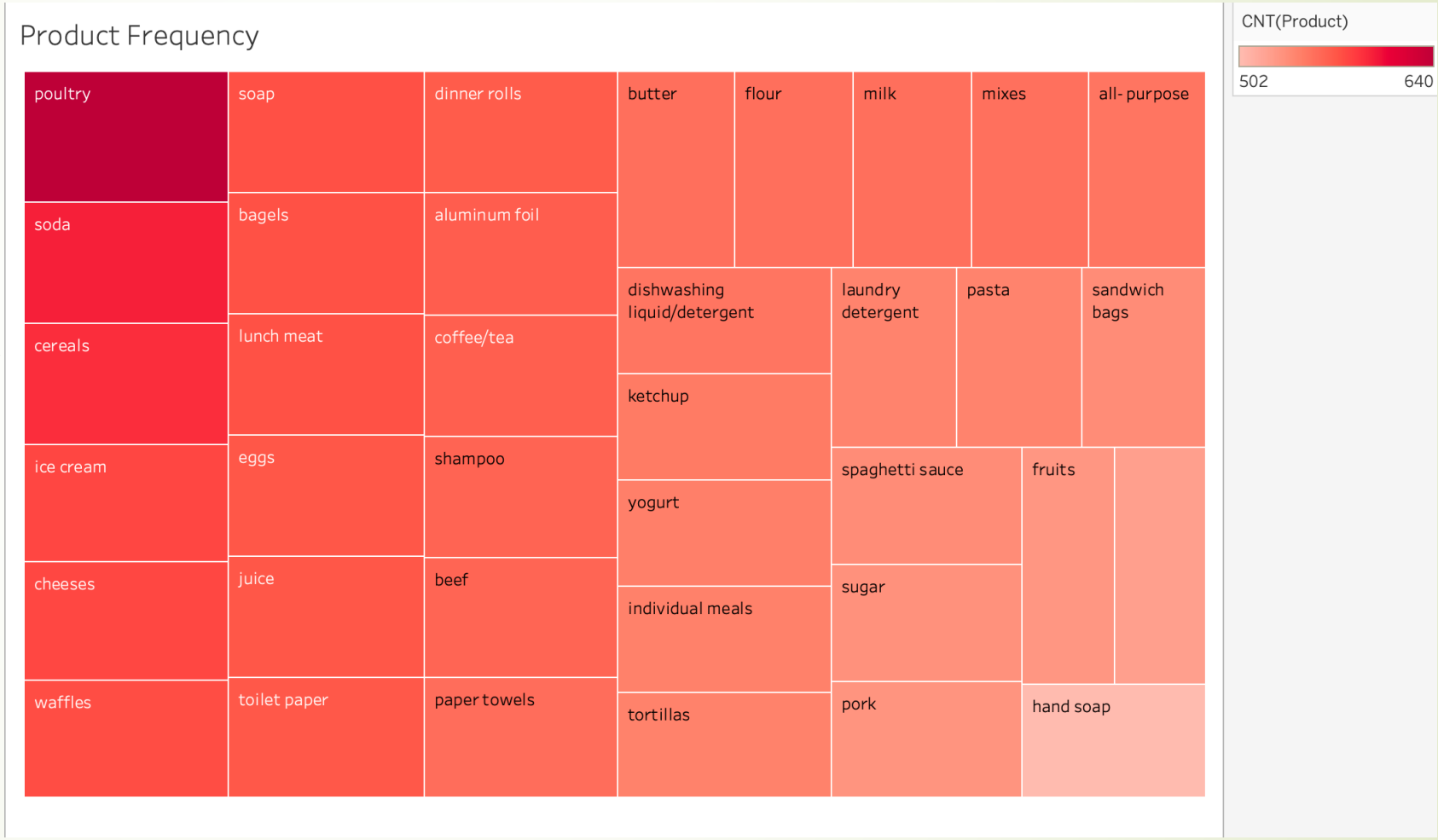
- To analyse the Point of Sale (POS) data from a grocery store to identify the frequently occurring combinations of items in customer orders. Provide actionable recommendations for implementing popular combo offers and discounts, aiming to increase the store's revenue and enhance customer satisfaction.



Exploratory Analysis and Inferences

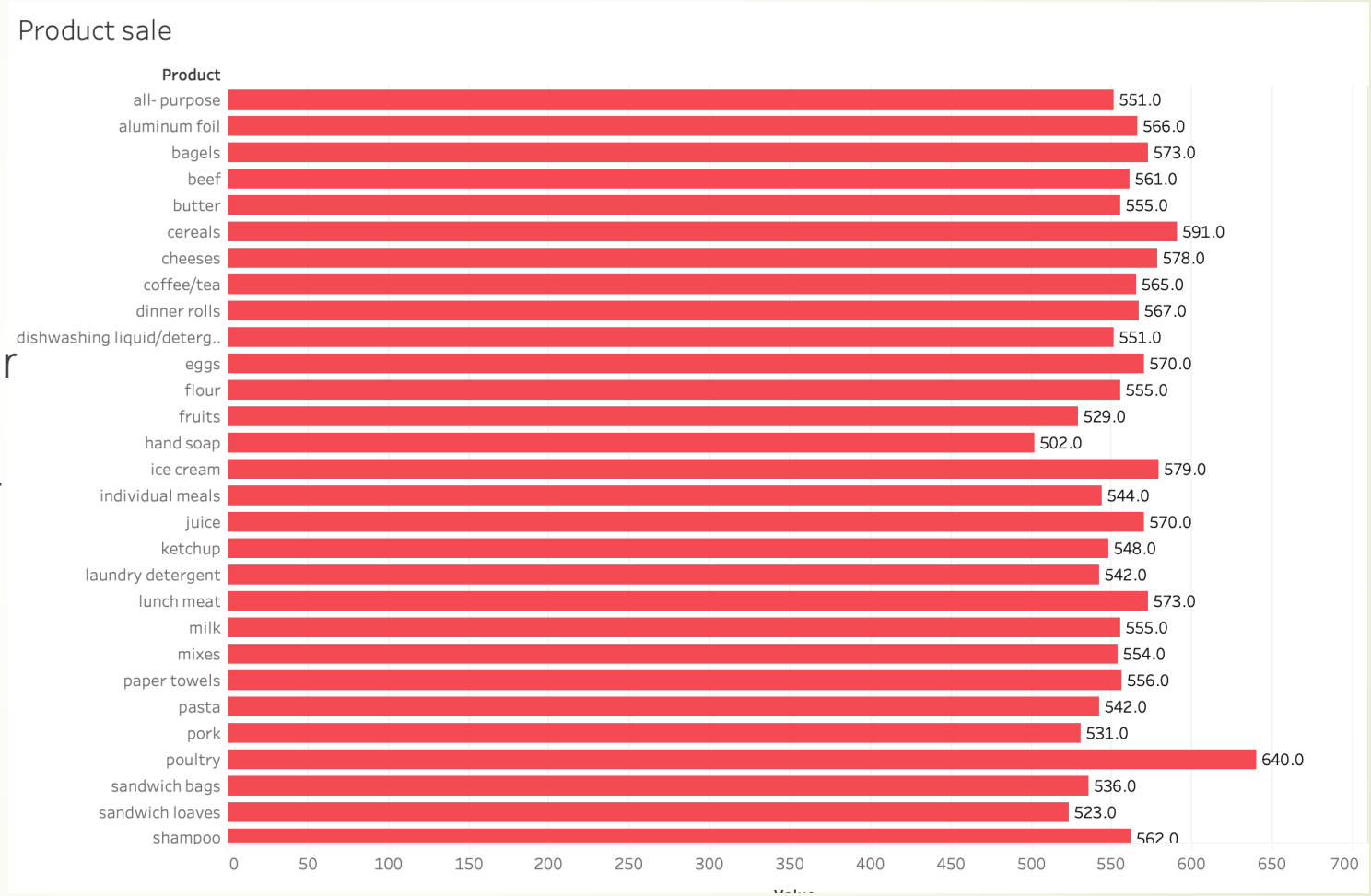
Frequency of products sold

- The tree map analysis reveals the topmost occurring products in the dataset, with Poultry having the highest count of 640, followed by Soda with a count of 597, Cereal with 591, Ice-cream with 579, Cheese with 578, and Waffles with 575. These findings highlight the popularity and demand for these specific products.



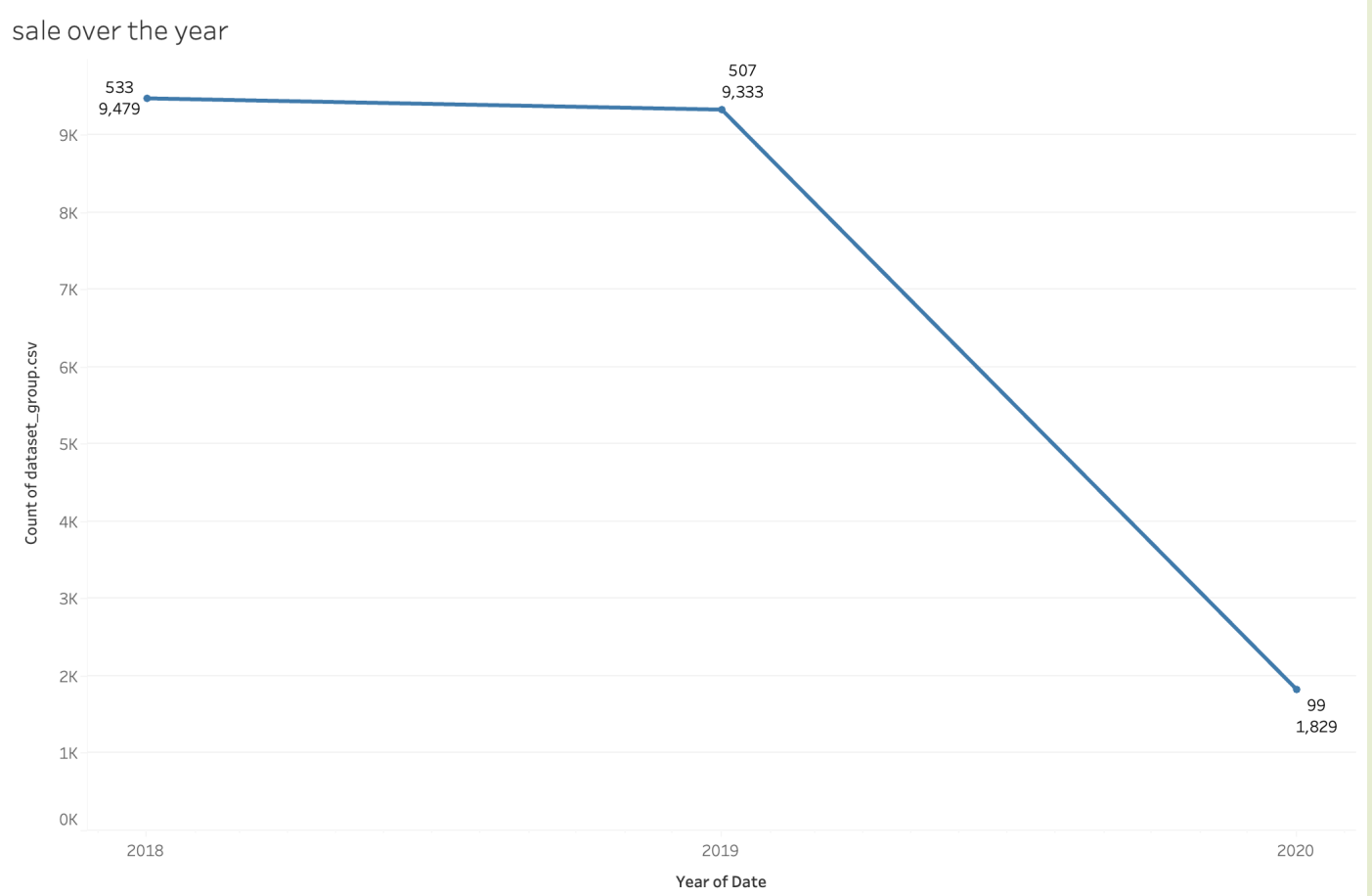
QUANTITY OF PRODUCT SOLD

- Poultry is the highest bought product followed by cereals
- The lowest sold items are hand soaps and sandwich loaves and flour
- The products are all equally bought with slight ups and downs



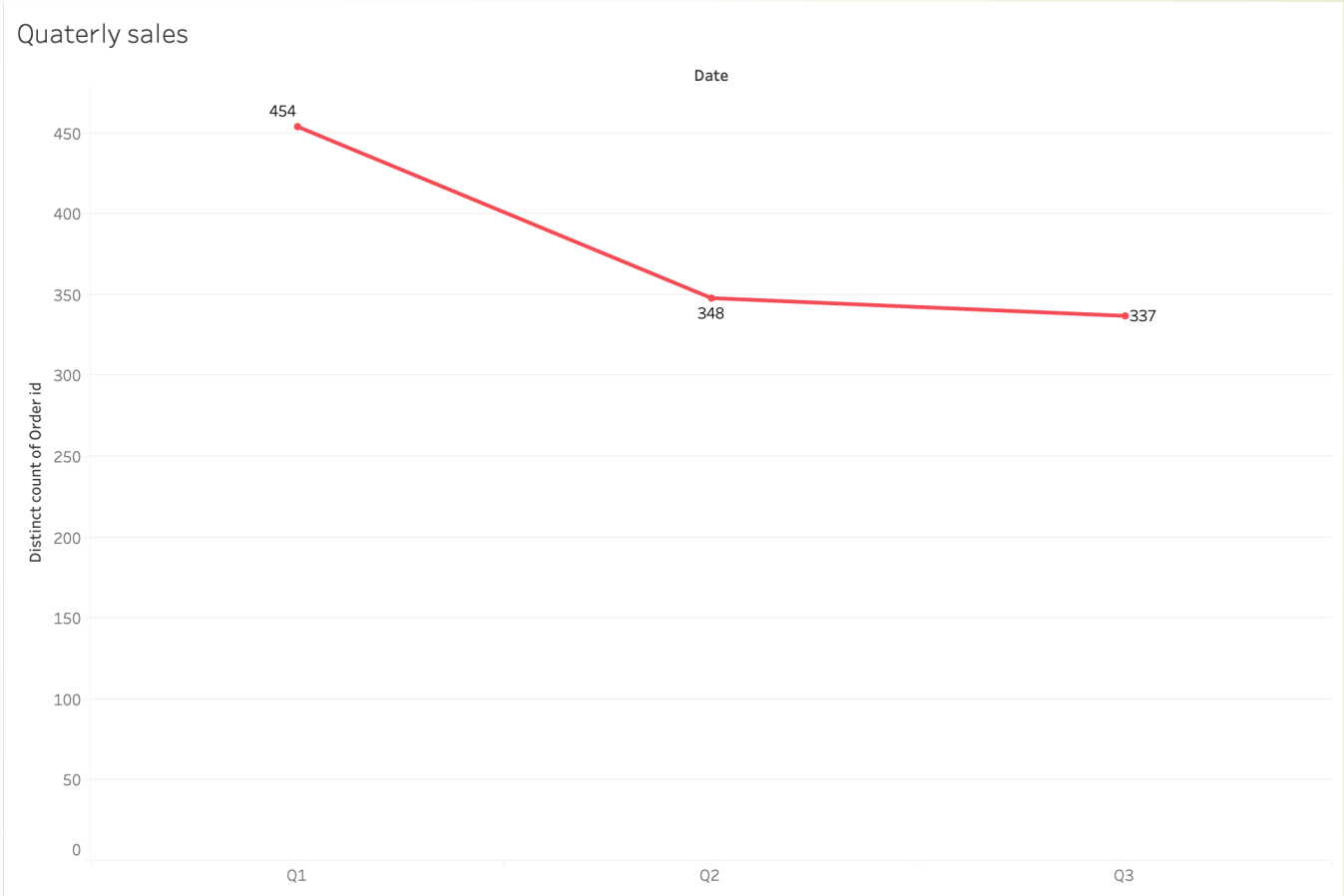
Yearly Sales

- Based on the below data, it can be inferred that there was a relatively stable level of order activity from 2018 to 2019, with a slight decrease observed. However, further analysis is required to understand the trends and patterns in the order data for 2020, since it is important to consider that the dataset only includes data for the first two months of 2020.



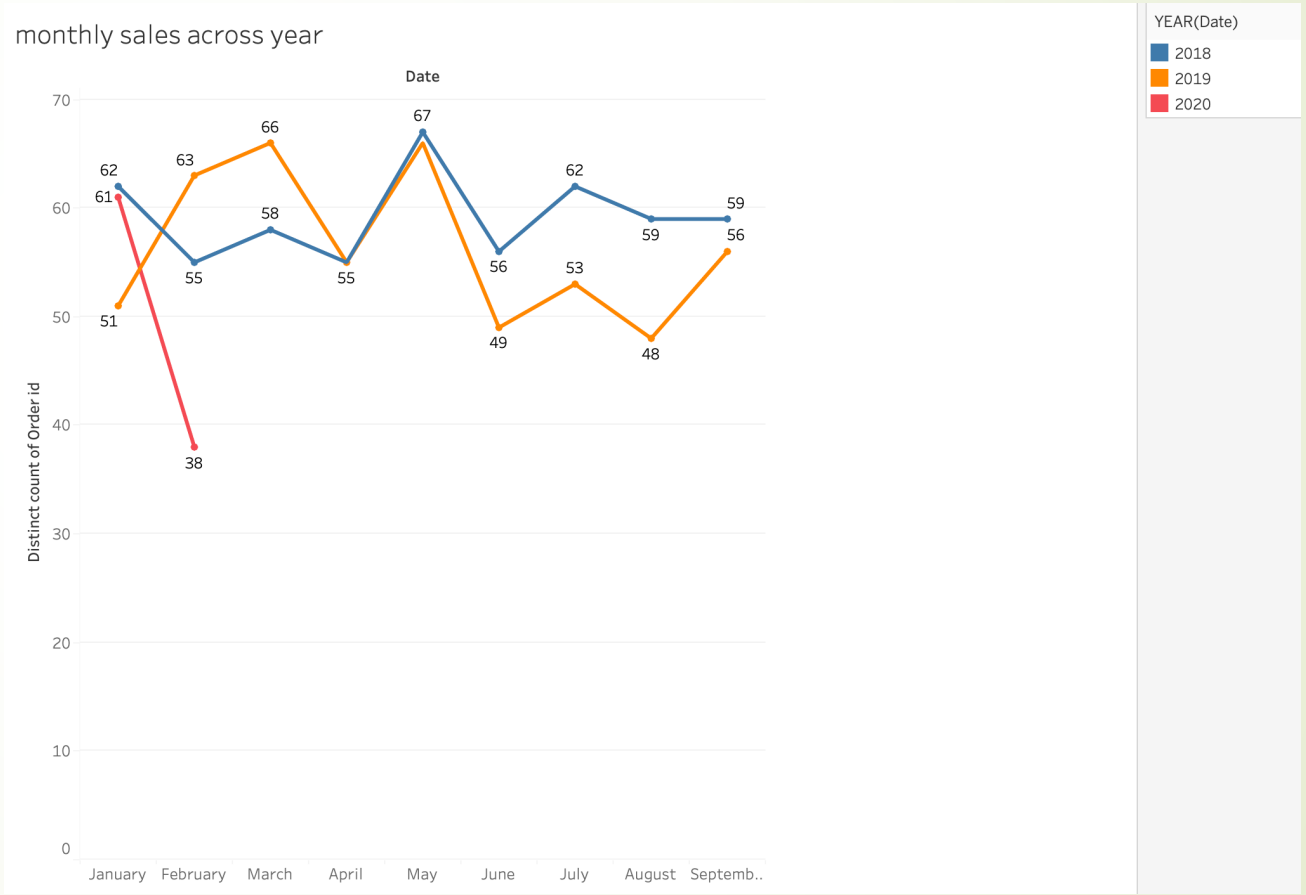
Quarterly Sales

- It can be inferred that there was a general downward trend in the number of unique order IDs from Q1 to Q3. This may indicate a potential decrease in overall order activity during this period. However, further analysis is required to understand the complete picture of order trends throughout the year, especially in the absence of Q4 data.



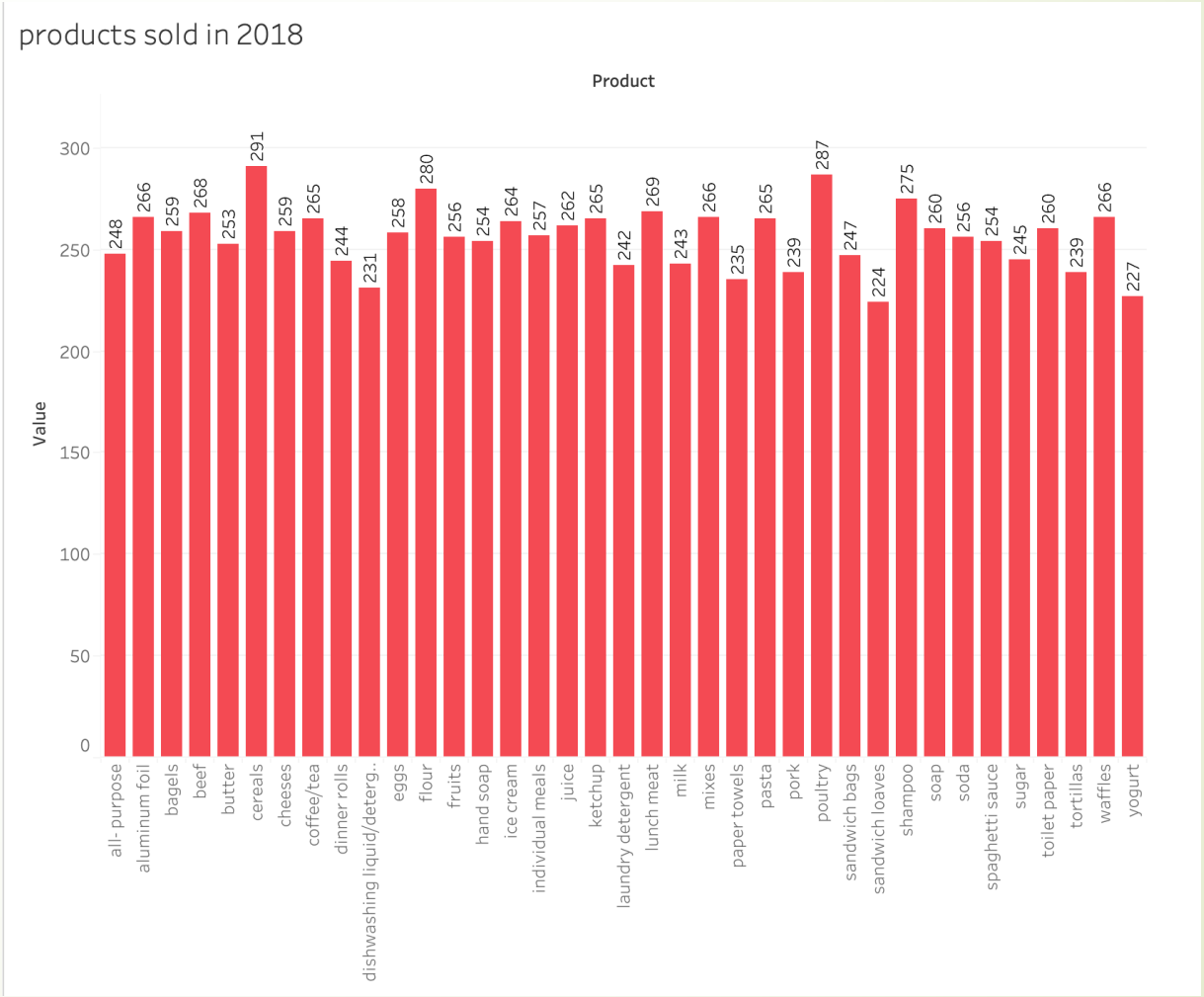
Monthly Orders

- The data indicates variations in customer ordering patterns throughout the year.
- The highest number of unique orders was observed in January, followed by February and May. June had the lowest count of unique orders. It is worth noting that there is a relatively consistent level of order activity during the summer months (June, July, and August) and a slight increase in September



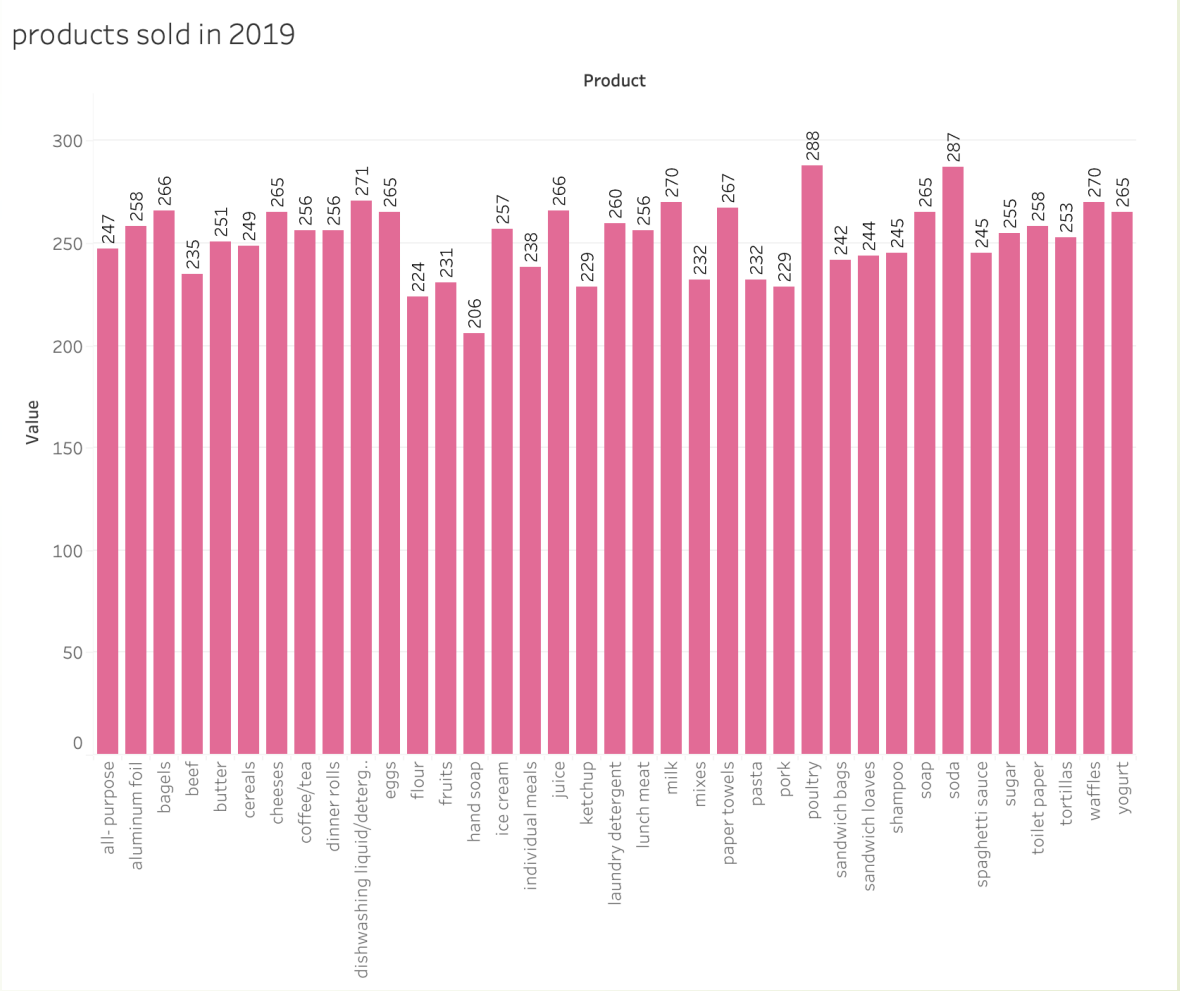
2018 – yearly data

- In 2018, the top five products that were most frequently ordered by customers were cereals, poultry, flour, shampoo, and lunch meat. Cereals ranked first with 291 orders, followed closely by poultry with 287 orders. Flour, shampoo, and lunch meat also had significant demand, with 280, 275, and 269 orders respectively. These products demonstrated their popularity and appeal to customers, indicating a strong market demand for these items during that year.



2019 – Yearly data

- In 2019, the top five products that were highly ordered by customers were poultry, soda, dishwashing liquid, waffles, and milk. Poultry took the lead with 288 orders, closely followed by soda with 287 orders. Dishwashing liquid, waffles, and milk also had a substantial number of orders, with 271, 270, and 270 respectively. These products showcased their popularity and customer preference, indicating a strong demand for them throughout the year.



2020 – Yearly data

- In 2020, the top five products that were highly ordered by customers were dinner rolls, poultry, pork, ice cream, and beef. Dinner rolls took the lead with 67 orders, closely followed by poultry with 65 orders. Both pork and ice cream were popular choices, each receiving 63 and 58 orders respectively. Beef also had a notable number of orders with 58. These products demonstrated their appeal and customer demand, showcasing their popularity of 2020.

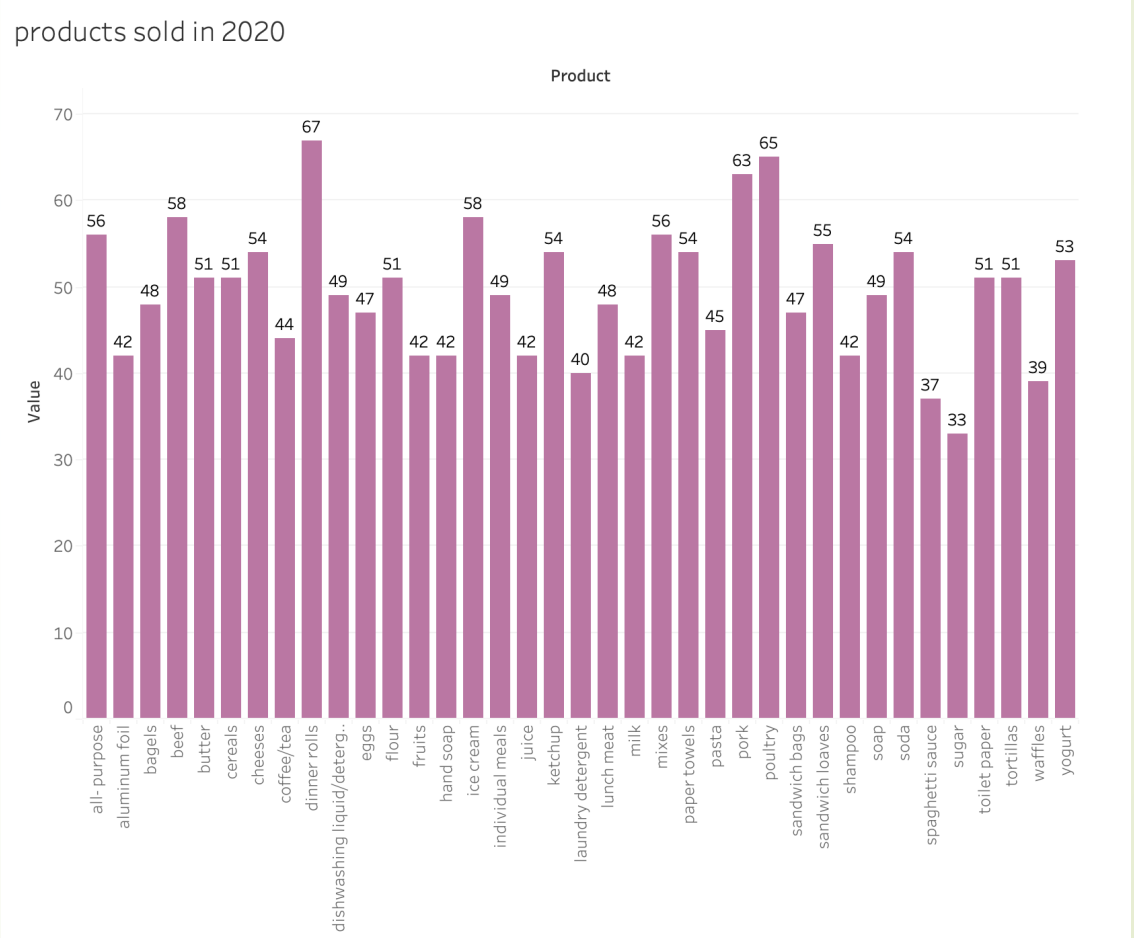


Tableau Link

- https://public.tableau.com/views/Apoorva_P_06_May_2024_MRA_problem2/productssoldin2020?:language=en-GB&:sid=&:display_count=n&:origin=viz_share_link

Market Basket Analysis

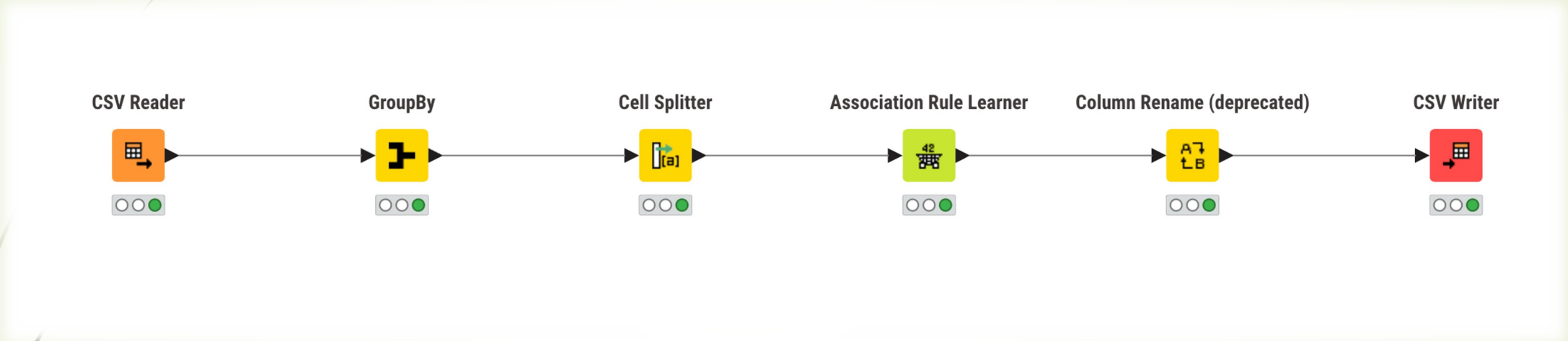
Market Basket Analysis

- Market Basket Analysis is a technique used to identify patterns and associations among products that are frequently purchased together in customer transactions. It helps businesses understand customer behaviour and uncover product relationships, which can be used to optimize sales strategies and increase revenue. By analysing transactional data, Market Basket Analysis generates insights on item co-occurrence and association rules, enabling businesses to make informed decisions on product bundling, cross-selling, and targeted marketing campaigns. This analysis provides valuable insights into customer preferences, allowing businesses to improve customer satisfaction and drive business growth.

Association Rules

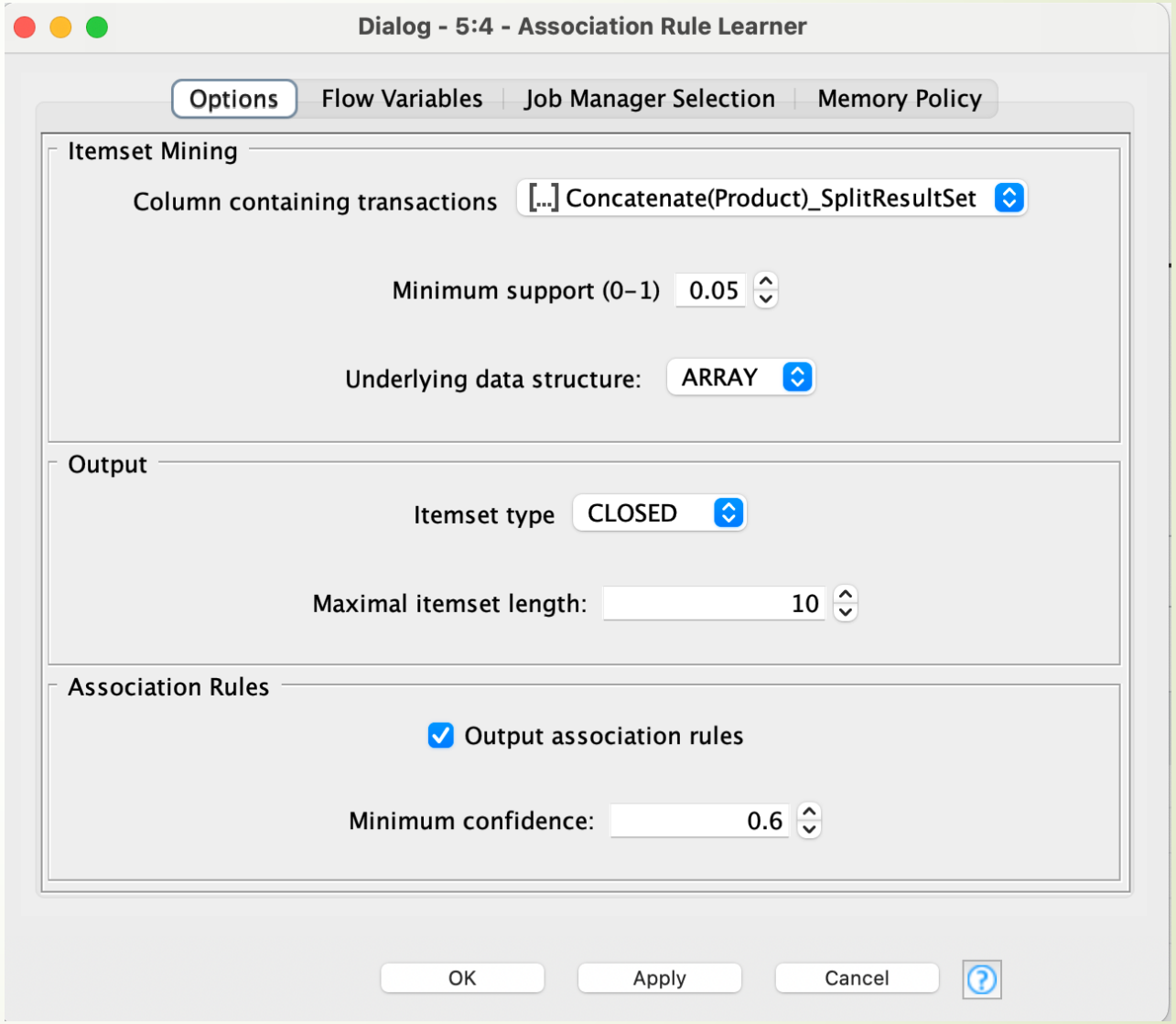
- Association rules in Market Basket Analysis reveal the relationships and co-occurrence patterns between items, providing valuable insights into customer purchasing behaviour and preferences.
- The relevance of association rules lies in their ability to guide businesses in optimizing product placement, creating targeted marketing campaigns, and implementing effective cross-selling and upselling strategies to enhance customer satisfaction and increase revenue.

KNIME workflow image



Threshold values of Support and Confidence

- Threshold value for Minimum Support is 0.05
- Threshold value for Minimum Confidence is 0.6
- In this analysis, we have defined the threshold values for Support and Confidence as 0.05 and 0.6, respectively. These values help us determine which association rules are significant and trustworthy. By using these thresholds, we can filter out less important rules and focus on the ones that have strong support and confidence levels, ensuring that our analysis provides meaningful insights for decision-making.



Associations Identification

OUTPUT DATA

MRA Project 2 -Apoorva_P

Support	Confidence	Lift	Recommended_item	Recommended_with	Items_list
0.05004389815627740	0.6404494382022470	1.700400722872630	juice	<---	[yogurt, toilet paper, aluminum foil]
0.05004389815627740	0.6195652173913040	1.644952873213740	juice	<---	[yogurt, poultry, aluminum foil]
0.05004389815627740	0.6129032258064520	1.6159647550776600	coffee/tea	<---	[yogurt, cheeses, cereals]
0.05004389815627740	0.6	1.4237500000000000	poultry	<---	[dishwashing liquid/detergent, laundry detergent, mixes]
0.05092186128182620	0.6304347826086960	1.6777224705404300	mixes	<---	[yogurt, poultry, aluminum foil]
0.05092186128182620	0.6105263157894740	1.6596407486496700	sandwich bags	<---	[cheeses, bagels, cereals]
0.05092186128182620	0.6744186046511630	1.7262085184217400	cheeses	<---	[bagels, cereals, sandwich bags]
0.05092186128182620	0.6170212765957450	1.5582865499834900	cereals	<---	[cheeses, bagels, sandwich bags]
0.05092186128182620	0.6304347826086960	1.6209147119442500	dinner rolls	<---	[spaghetti sauce, poultry, cereals]
0.05092186128182620	0.6373626373626370	1.5124084249084200	poultry	<---	[dinner rolls, spaghetti sauce, cereals]
0.05092186128182620	0.6041666666666670	1.5892513471901500	milk	<---	[poultry, laundry detergent, cereals]
0.05179982440737490	0.6276595744680850	1.6101447191872700	eggs	<---	[dinner rolls, poultry, soda]
0.05179982440737490	0.641304347826087	1.6488615173226000	dinner rolls	<---	[spaghetti sauce, poultry, ice cream]
0.05179982440737490	0.686046511627907	1.6279312015503900	poultry	<---	[dinner rolls, spaghetti sauce, ice cream]
0.05179982440737490	0.6276595744680850	1.613779357379570	dinner rolls	<---	[spaghetti sauce, poultry, juice]
0.05179982440737490	0.6020408163265310	1.42859268707483	poultry	<---	[dinner rolls, spaghetti sauce, juice]
0.05179982440737490	0.6344086021505380	1.627458103264560	eggs	<---	[paper towels, dinner rolls, pasta]
0.05179982440737490	0.6020408163265310	1.6210980846239200	pasta	<---	[paper towels, eggs, dinner rolls]
0.05355575065847230	0.6421052631578950	1.6509207556136400	dinner rolls	<---	[spaghetti sauce, poultry, laundry detergent]
0.05355575065847230	0.6559139784946240	1.556429211469530	poultry	<---	[dinner rolls, spaghetti sauce, laundry detergent]
0.05531167690956980	0.6237623762376240	1.564901644349460	ice cream	<---	[paper towels, eggs, pasta]
0.05531167690956980	0.63	1.6161486486486500	eggs	<---	[paper towels, ice cream, pasta]
0.05531167690956980	0.6428571428571430	1.731003039513680	pasta	<---	[paper towels, eggs, ice cream]
0.05531167690956980	0.6494845360824740	1.7911934300192200	paper towels	<---	[eggs, ice cream, pasta]

Support, Confidence, and Lift Values: Metrics for Association Analysis

Support: The support value represents the frequency or popularity of an itemset in the dataset. It indicates how often a specific combination of items appears together in customer transactions.

Confidence: Confidence measures the likelihood that a customer who buys one item will also purchase another item. It is calculated as the ratio of the number of transactions where both items are purchased together to the number of transactions where the first item is purchased.

Lift: Lift measures the strength of association between two items in an association rule. It compares the probability of the two items being purchased together to the probability of them being purchased independently. A lift value greater than 1 suggests a positive association, indicating that the items are more likely to be purchased together.

Suggestion of Possible Combos with Lucrative Offers

Recommendations

- **Combo Deal:** Offer a special combo deal where customers can buy yogurt, poultry, and aluminium foil along with juice to avail a discounted price or additional item.
- **Buy Two Get One Free:** Introduce a "buy two get one free" offer on dinner rolls, spaghetti sauce, and ice cream to incentivize customers to purchase these items together.
- **Bundle Promotion:** Create a bundle promotion where customers can buy paper towels, eggs, and pasta together at a discounted.
- **Cross-Selling Offer:** Provide a cross-selling offer where customers purchasing cereals can get a discount on cheese, bagels, and sandwich bags.
- **Limited-Time Promotion:** Launch a limited-time promotion where customers buying poultry, laundry detergent, and mixes can receive a percentage savings.
- **Loyalty Program:** Implement a loyalty program where customers who frequently purchase recommended items or participate in the suggested combos can earn rewards or exclusive discounts. This will incentivize customer loyalty and encourage them to continue shopping with the store, fostering long-term relationships and repeat purchases.

These recommendations are based on the association rules and the occurrence of certain items together, aiming to increase customer satisfaction and encourage them to explore additional products.


Thank
you