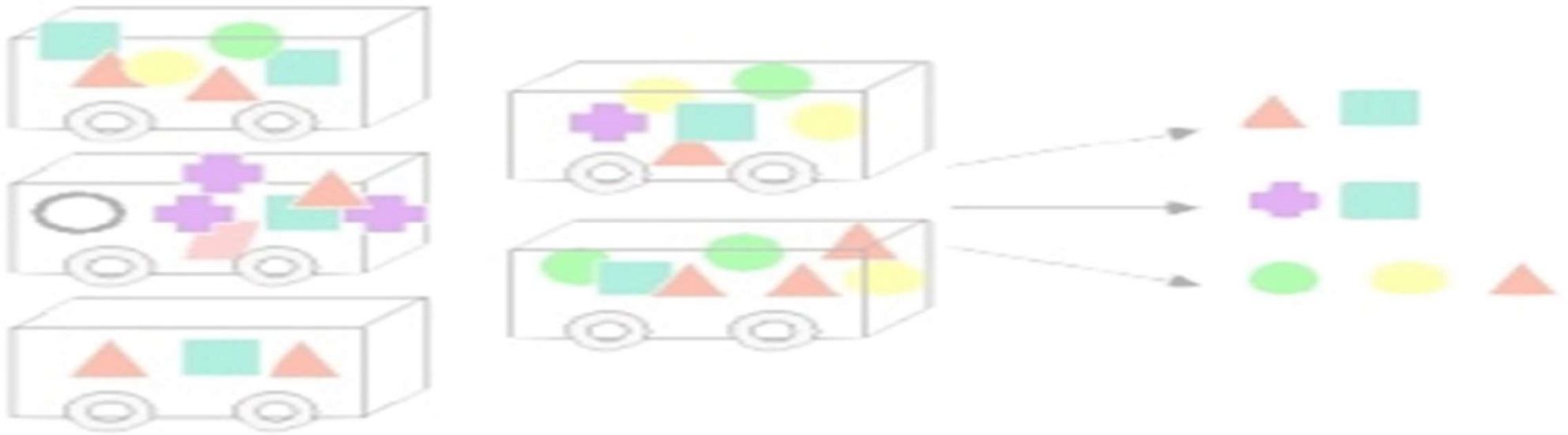


GROCERY STORE DATA



MARKET BASKET ANALYSIS

VIDYA V
PGPDSBA.O.2023 B

AGENDA



Analyze the grocery POS data to derive insights and identify patterns

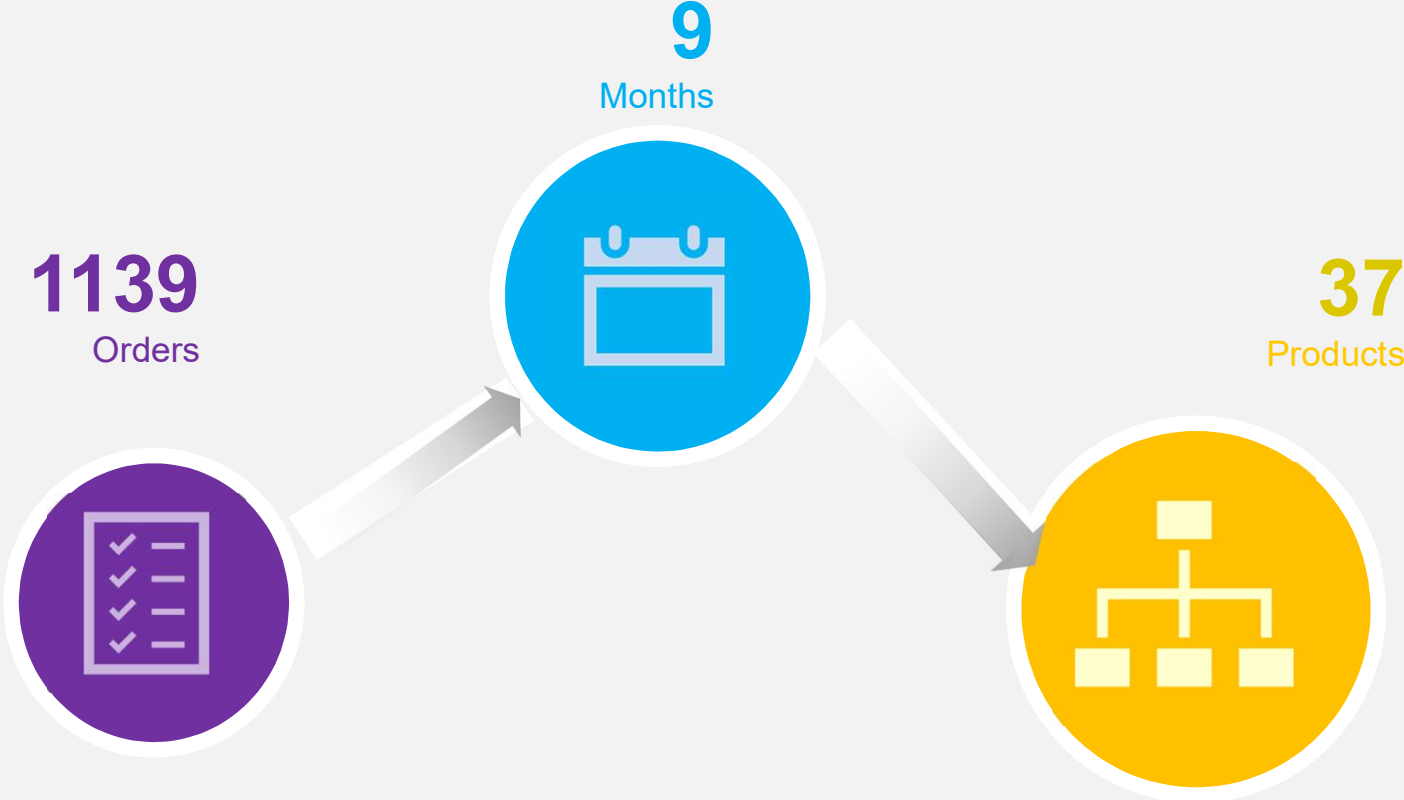


Perform Market basket analysis to understand buying patterns



Develop marketing strategies and plan offers

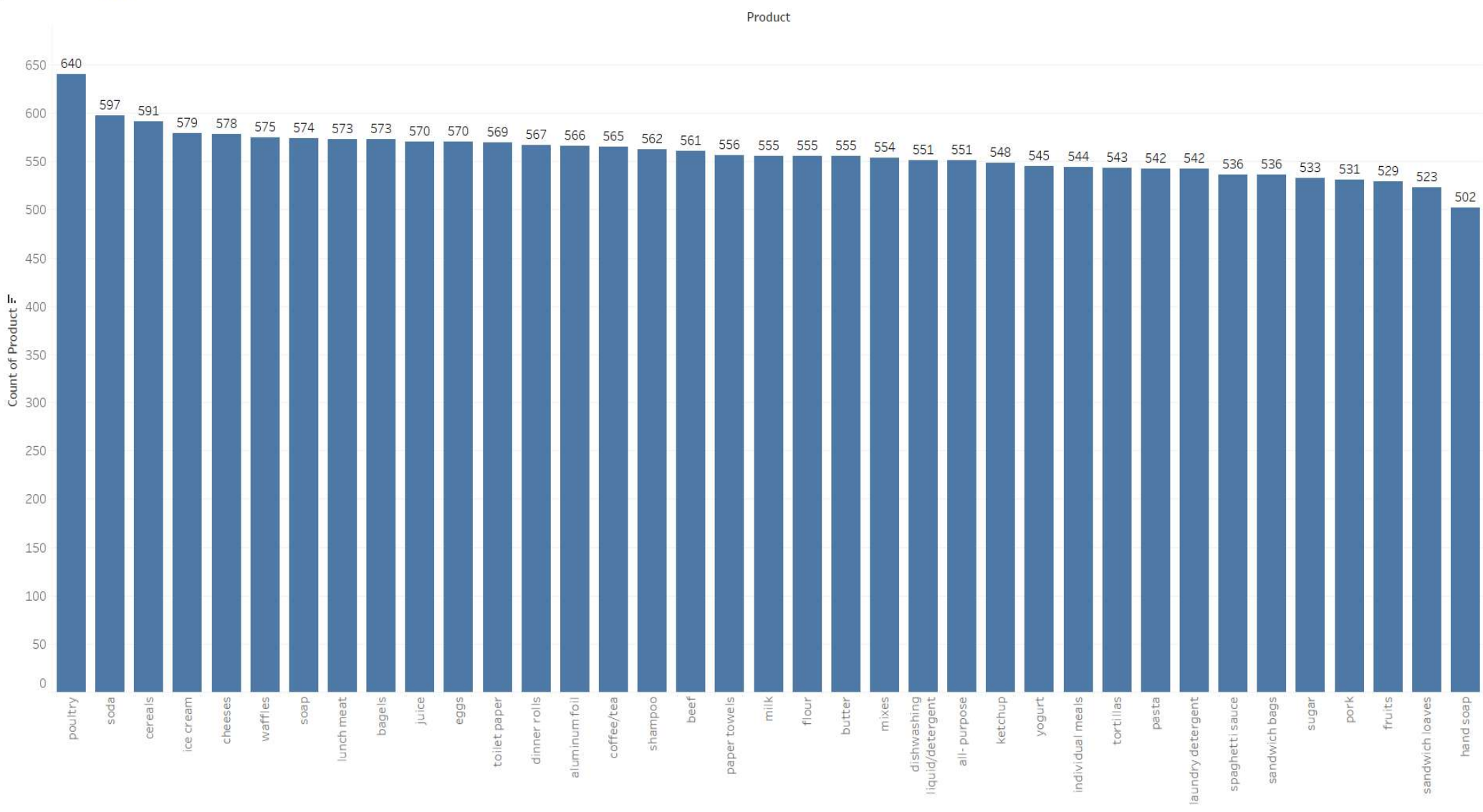
EXECUTIVE SUMMARY



A magnifying glass with a gold-colored frame and a dark red handle is positioned over a line graph. The graph features a blue line that trends upwards from left to right, with some minor fluctuations. The background of the graph consists of numerous thin, vertical gold-colored bars. The text 'UNIVARIATE ANALYSIS' is centered within the magnifying glass's lens in a bold, dark blue, sans-serif font. On the right side of the image, a vertical axis is marked with numerical values: 10, 20, 30, 40, and 50.

UNIVARIATE ANALYSIS

product counts



A magnifying glass with a gold-colored frame and a dark red handle is positioned over a chart. The chart features a blue line graph and yellow vertical bars. The text "BIVARIATE ANALYSIS" is centered within the magnifying glass's lens. The background is a light blue gradient.

BIVARIATE ANALYSIS

50

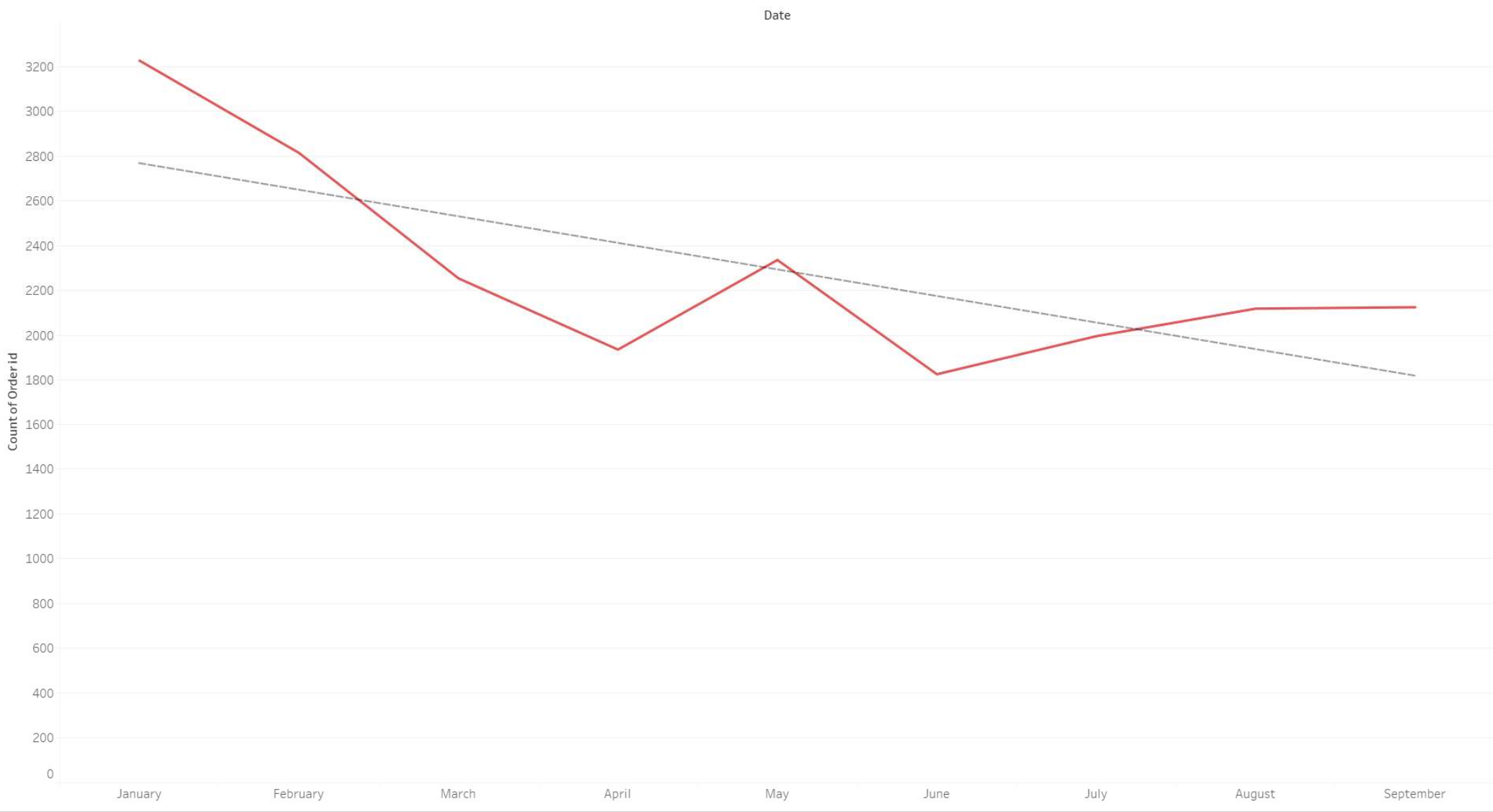
40

30

20

10

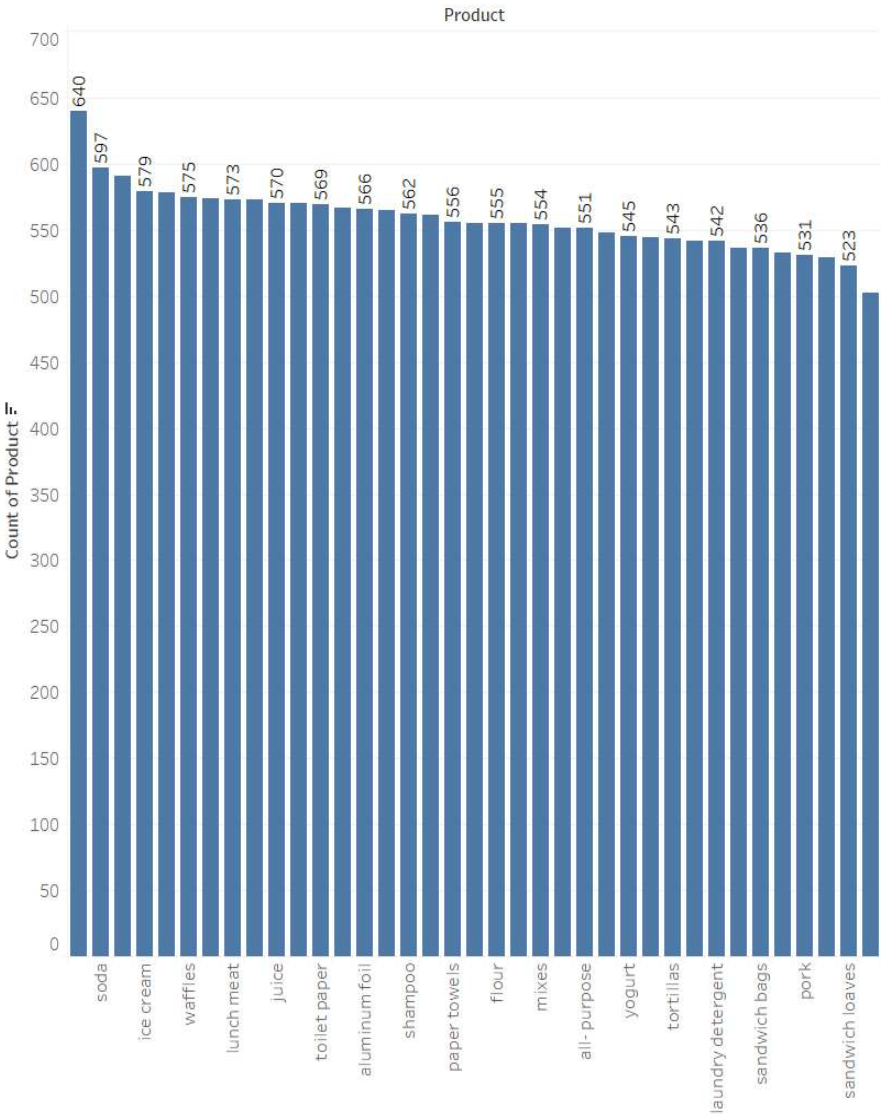
orders trend



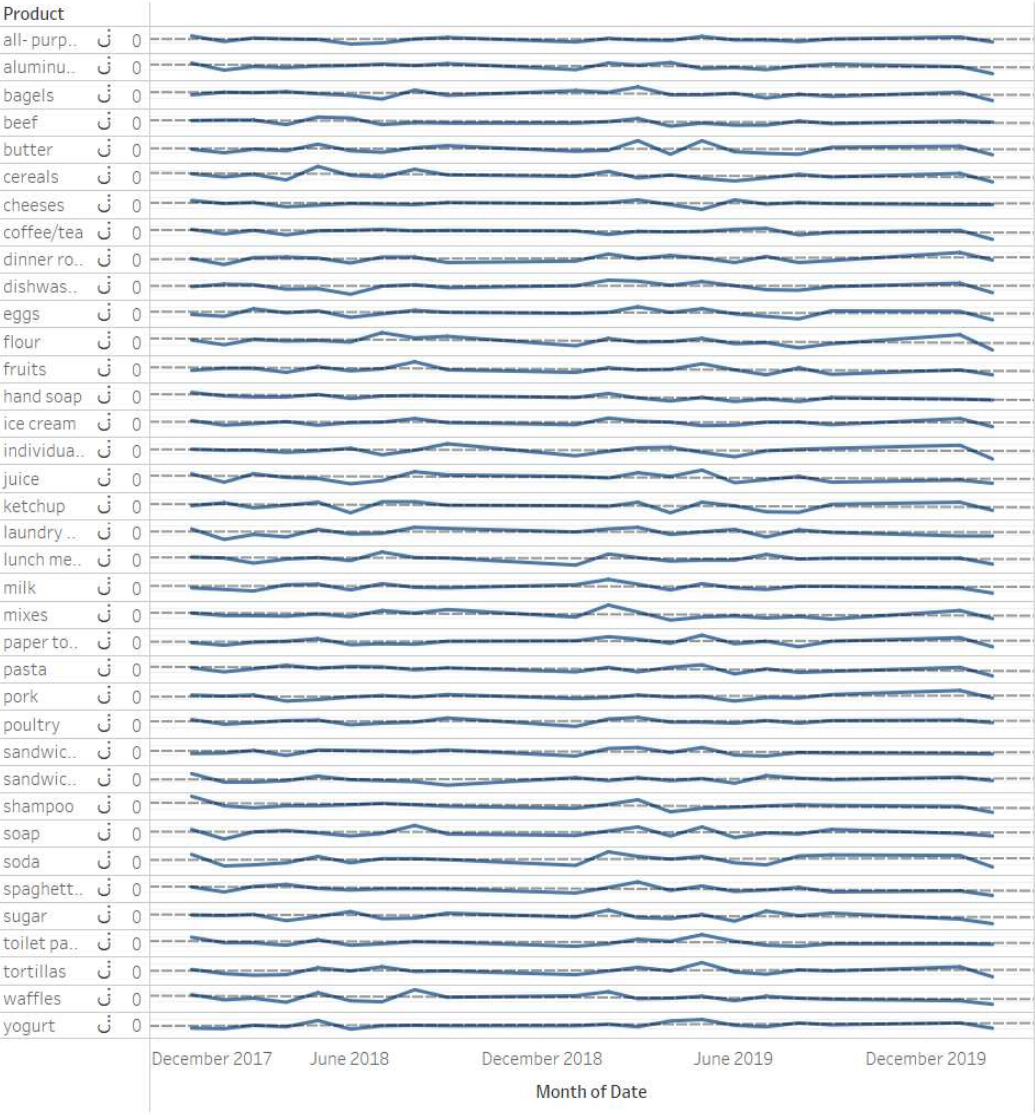
A magnifying glass with a gold-colored frame and a dark red handle is positioned over a line graph. The graph features a blue line that trends upwards from left to right, with some minor fluctuations. The background of the graph consists of numerous thin, vertical gold bars of varying heights. The magnifying glass's lens is centered on the text 'MULTIVARIATE ANALYSIS', which is written in a bold, dark blue, sans-serif font. The overall image conveys a sense of detailed examination and data analysis.

MULTIVARIATE ANALYSIS

product counts



oc-prod-ts



EDA OBSERVATIONS

- Poultry is the most bought item- with a total purchase count of 640
- Hand soap is the least bought item- with a purchase count of 502
- The trend of order counts is decreasing- which might be a cause for concern
- On individually inspecting the purchase trend of each product, for most of the products, the lowest purchase count was in June
- After which there was a general increasing trend, followed by another decline in August
- After August, most of the products have remained stagnant or have declined
- Of the total 37 products, poultry, soda, dinner rolls, paper towels, yogurt, tortillas, pork, sandwich loaves are the only ones that show a positive increasing purchase trend
- But none of the products have reached the purchase counts of that in January



MARKET BASKET ANALYSIS

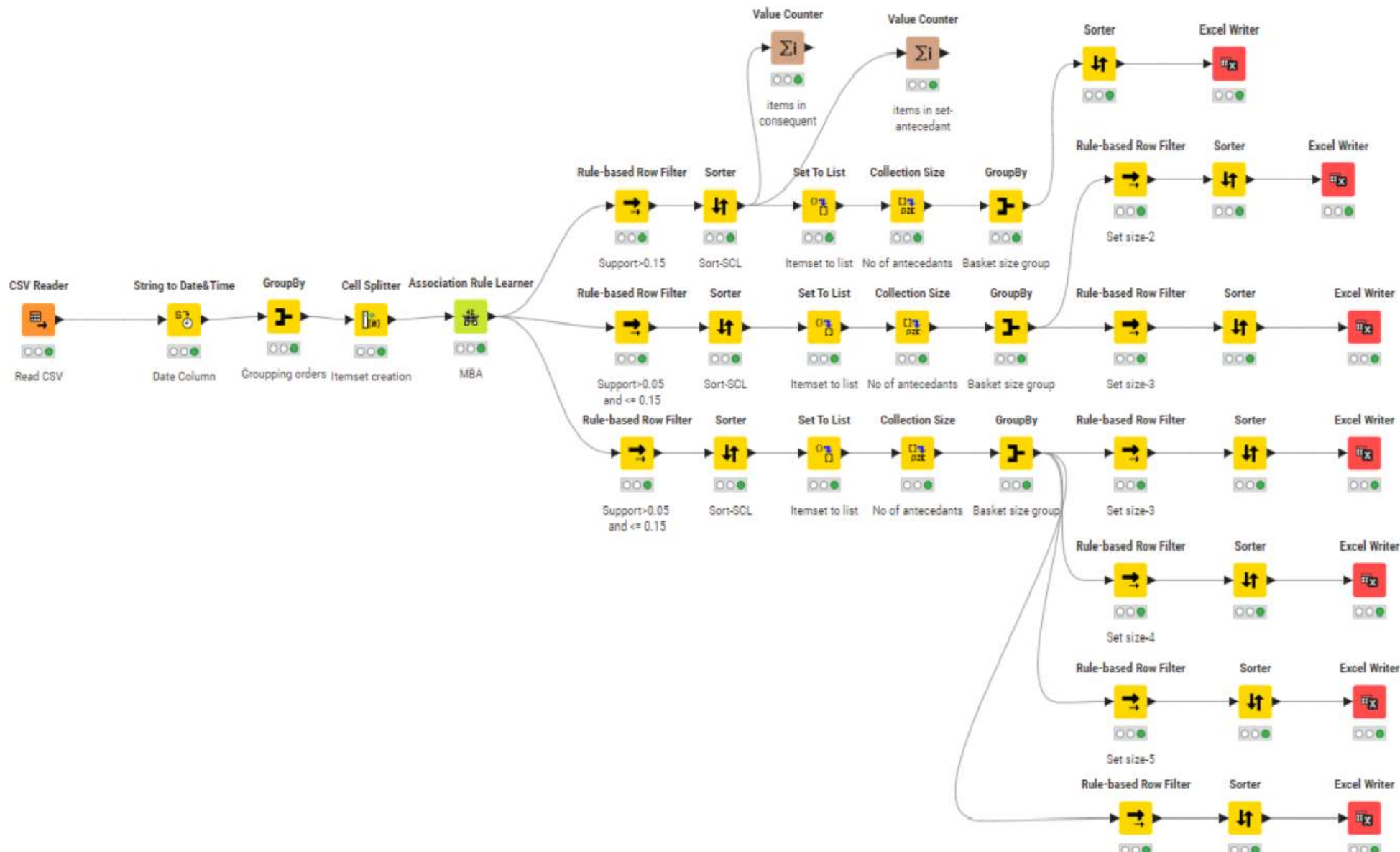
MARKET BASKET ANALYSIS

- Technique used identify associations between frequently bought products
- Objective: To identify associations between the purchase of products to improve sales, product placements and develop marketing strategies
- This is done by the association rule learner that uses various algorithms to identify the associations between products
- Important metrics of association rule learner:
 - Support: Measure of the popularity of the antecedent. Ranges from 0-1. High support indicates a high probability of the antecedents. For the purpose of this project, the support value is assumed to be 0.01
 - Confidence: Measure of association between the consequent and the antecedent. Higher, the confidence more likely the consequent to be purchased. Assumed confidence value- 0.45
 - Lift: Measure of the strength of association between the antecedent and consequent

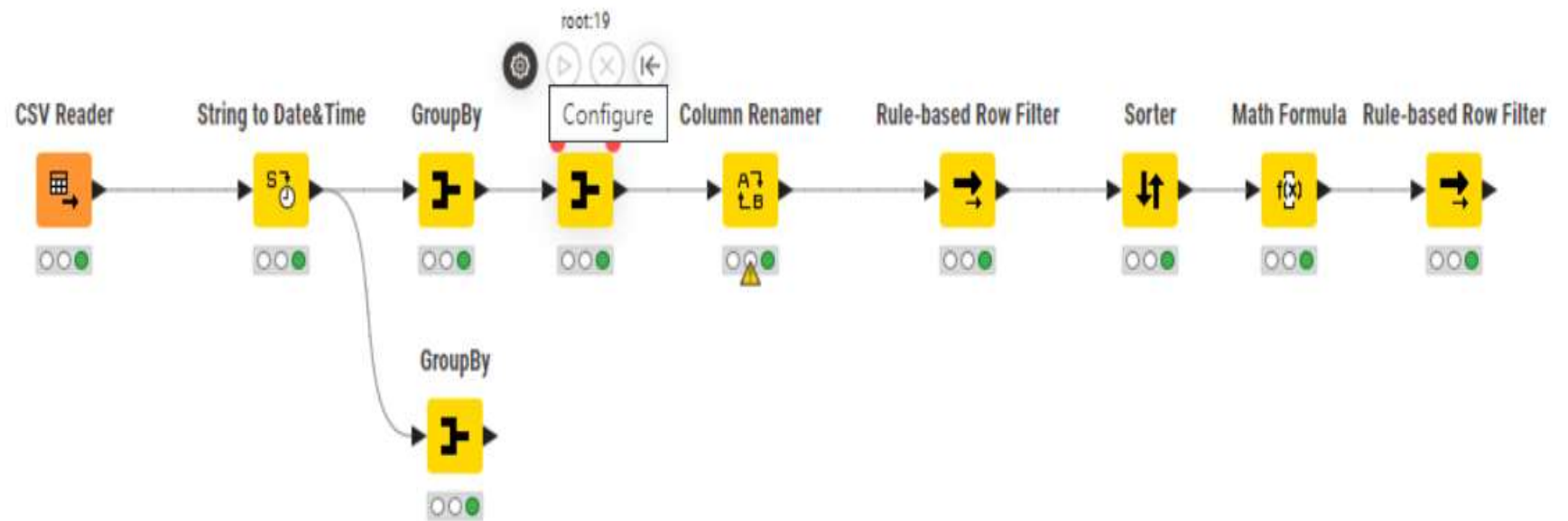
MARKET BASKET ANALYSIS

- In the dataset given, there were duplicate values in the itemsets. In order to optimize the analysis, a separate kind of analysis was done to identify items that are frequently bought in multiples
- Thus, two kinds of analysis was done:
 - To identify association between different products and to enable the suggestion of combo offers – association rule learner and filters
 - To identify the association of the product with itself in order to explore options of buy 1 get 1 kind of offers- groupby order and product and identify the mode of occurrences

KNIME WORKFLOW- ASSOCIATION RULE LEARNER



KNIME WORKFLOW- IDENTIFICATION OF SELF- ASSOCIATION



OBSERVATIONS- Associations

Itemset Size	Antecedants	Consequent	Support	Confidence	Lift
1	Dinner Rolls	Poultry	0.195	0.501	1.189
1	Poultry	Dinner Rolls	0.195	0.462	1.189
1	Dishwashing Liquid/Detergent	Poultry	0.187	0.482	1.144
1	Eggs	Soda	0.187	0.48	1.228
1	Soda	Eggs	0.187	0.479	1.228
2	Spagetti Sauce, Poultry	Dinner Rolls	0.099	0.579	1.49
2	Spagetti Sauce, Dinner Rolls	Poultry	0.099	0.577	1.368
2	Dinner Rolls, Poultry	Spagetti Sauce	0.099	0.509	1.364
3	Eggs, ice Cream,pasta	paper towels	0.055	0.649	1.791
3	Eggs, ice Cream,paper towels	pasta	0.055	0.643	1.731
3	Paper towels, ice Cream,pasta	eggs	0.055	0.630	1.616
3	Eggs, paper towels,pasta	ice cream	0.055	0.624	1.565
4	poultry, fruits, toilet paper, sugar	cereals	0.027	0.795	2.120
4	cheeses, bagels, coffee/tea, sandwich bags	poultry	0.027	0.795	2.007
4	dinner rolls, hand soap, juice, sugar	dinner rolls	0.027	0.795	1.886
4	eggs, hand soap, pasta, individual meals	cheeses	0.027	0.775	1.993
4	bagels, cereals, coffee/tea, sandwich bags	paper towels	0.027	0.775	1.984

OBSERVATIONS- Self associations

Product	usual purchase count	No of orders for the purchase count	Percentage of Total orders
flour	2	105	9.218613
cereals	2	99	8.691835
juice	2	99	8.691835
poultry	2	98	8.604039
soap	2	98	8.604039

INFERENCES

- The associations of itemset and consequents are tabulated in the previous slide
- The more the number of items in set, less is the support
- Dinner rolls and poultry has the highest support of 0.195 and poultry has the highest confidence as well
- Juice, paper towels and yogurt are the most common antecedents in sets having support greater than 0.15
- Poultry is the most common consequent in sets with support > 0.15
- Flour is the product that is bought in multiple quantities most of the times

RECOMMENDATIONS



For items with high support and confidence, like dinner rolls and poultry, bundling them and selling them as a combo would help boost sales



For products like flour, cereals etc. that are mostly bought in multiple quantities in an order, buy 1 get 1 free offers can be considered



For items like hand soap, which have a very low purchase count, offering or increasing discounts is suggested

THANK YOU