MRA Project
Part 2

HIYA SHAH

GROCERY STORE MARKET BASKET ANALYSIS

hiyashah2020@gmail.com

TABLE OF CONTENT

- Problem statement

- Executive summary

- Objective

- EDA

- Inferences

- Market Basket Analysis

- KNIME workflow

- Recommendations

- Offers

PROBLEM STATEMENT

- In the highly competitive grocery retail industry, understanding customer buying patterns is crucial for enhancing sales, increasing customer satisfaction, and improving profitability. By identifying frequently purchased item combinations, grocery stores can craft effective marketing strategies, optimize inventory management, and tailor promotions to meet customer needs. Leveraging Point of Sale (POS) data can unlock valuable insights that drive customer-centric offerings, such as combo packs, discounts, and targeted promotions, which can increase basket size and improve customer retention. This analysis aligns with business goals by maximizing revenue, reducing operational costs, and boosting customer loyalty.

EXECUTIVE SUMMARY

- We have received the 2 years and 2 months data of a Grocery store. Consisting 20641 entries with 3 variable details regarding the demography of the transaction and item information.

OBJECTIVE

As a business analyst, the goal is to analyze the POS transactional data to identify frequently purchased item combinations. Using association rule mining or similar techniques, the aim is to uncover patterns that will help the store create targeted combo offers and discounts, ultimately driving revenue growth by increasing customer purchases and average basket size.

The project involves conducting a thorough analysis of Point of Sale (POS) Data for providing recommendations through which a grocery store can increase its revenue by popular combo offers & discounts for customers.

DATA SUMMARY

▪ A Grocery Store shared the transactional data with you. Your job is to identify the most popular combos that can be suggested to the Grocery Store chain after a thorough analysis of the most commonly occurring sets of menu items in the customer orders. The Store doesn't have any combo meals. Can you suggest the best combo meals?

▪ Below is head of the dataset

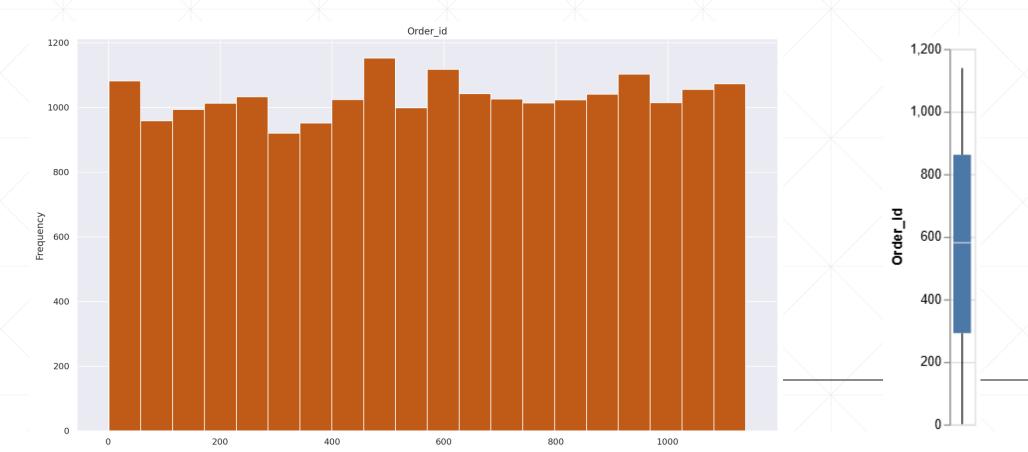|	Date	Order_id	Product
0	2018-01-01	1	yogurt
1	2018-01-01	1	pork
2	2018-01-01	1	sandwich bags
3	2018-01-01	1	lunch meat
4	2018-01-01	1	all- purpose

ABOUT DATA

• There are 20641 rows and 3 columns.

• There are 1 Integer and 2 object data types.

• There are no null values in the dataset.

• There are 4730 duplicate rows in the dataset.

• Data is available for 3 years 2018, 2019 and 2020

• Data for each year 2018 and 2019 is provided until Q3 ie; September only and for 2020 it is provided for January and February only

• A total of 1139 orders were received during this timeline.

• There are 37 unique Products as of today with the store.

Exploratary Data Analysis

UNIVARIATE ANALYSIS
DATA DISTRIBUTION & OUTLIERS

- There are no outliers, no high skewness in Order_id and the data seems to be normally distributed as well.
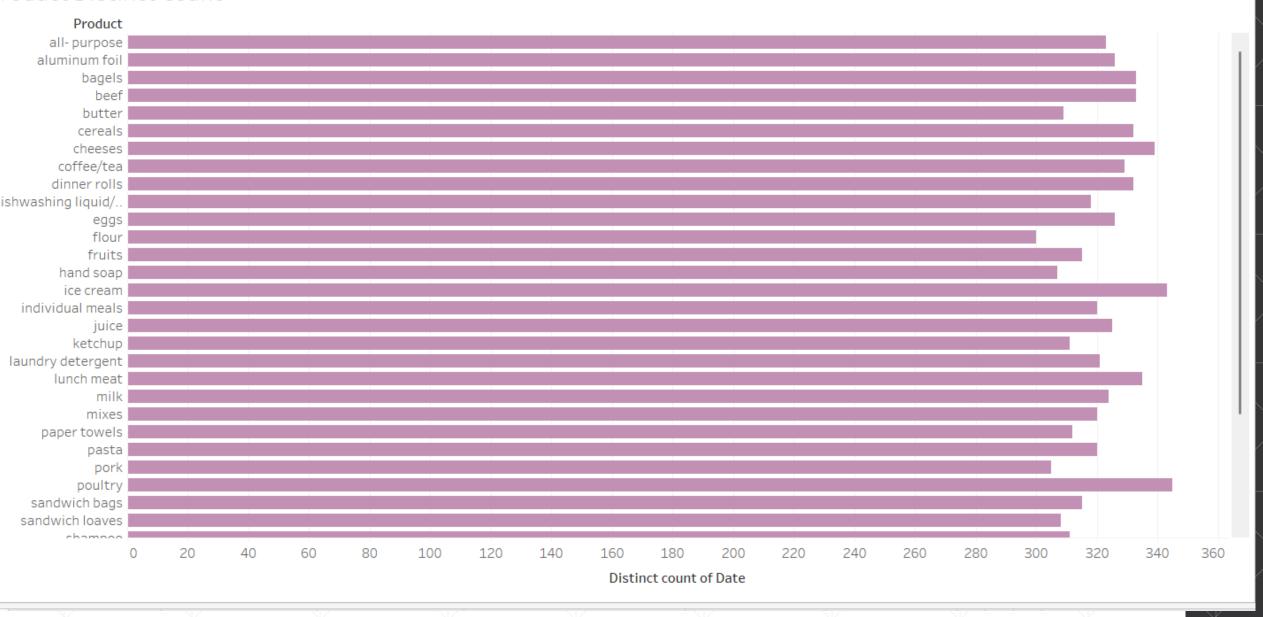
PRODUCT QUANTITY ORDERED MOST

Total Product Count
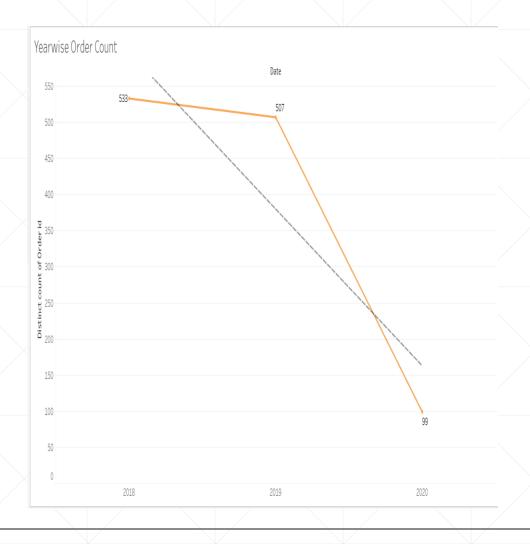
- Poultry accounts for highest purchased product quantity followed by Soda in 2nd and Cereal in the 3rd position.
- Hand Soup is the least purchased product quantity followed by Sandwich loaves and Fruits.
- There is not a drastic difference between the range in which all the products quantity are ordered.
- The milk, soap, coffee/tea, soda, cheese are more or less holds the same amount of orders.

Product Distinct Count

Product

Distinct count of Date

YEARLY ORDERS AND TREND

Yearwise Order Count
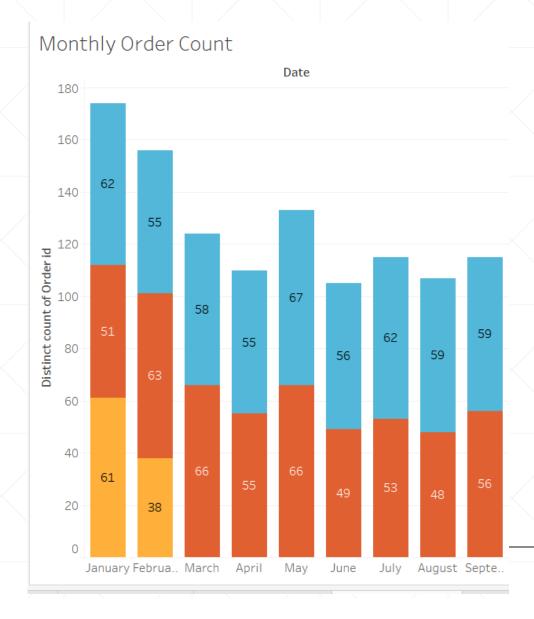
- The total number of orders are highest in the year 2018.

- Followed by 507 orders in 2019.

- 2020 has registered only 99 orders. As only two months data is given in the data set , this might be a reason for low order count.

- Trends shows that the number of orders placed has been decreasing yearly.

- Also it shows a downward trend and no forecast has been generated with this data.

- The

-      R-Squared value is- 0.79476

-      P-Value – 0.299317

QUARTERLY ORDERS AND TREND

- In 2018, Q3 had the highest number of orders with 180.

- In 2019, Q1 had the highest with 180.

- This started decreasing as quarters passed by. The lowest was recorded in Q1 of 2020.

- Also, it is evident that Q4 of every year doesn't have any sales.

- No proper trend can be analysed in terms of Quarter sales.

- 2018-
    - R-squared – 0.986842
    - P- Value – 0.0731864

- 2019-
    - R-squared – 0.994361
    - P- Value – 0.0478513

- 2020-
    - R-squared – 1
    - P- Value – N/A

Quarterly Order Count

MONTHLY ORDERS

Monthly Order Count
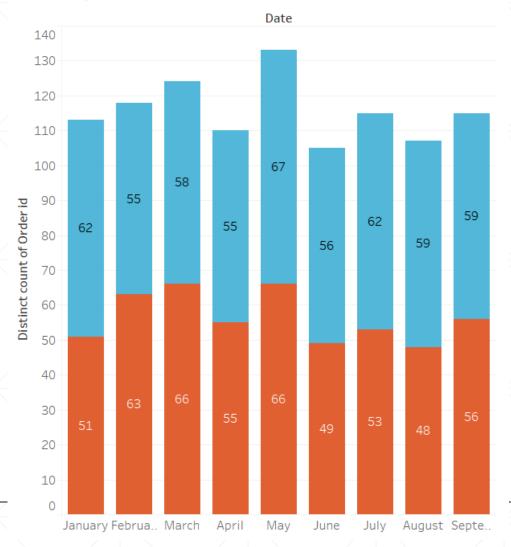
- There is no trend and seasonality available in the data provided.

- As per the forecast the sales would tend to increase in the mid quarters of the years.

- January, February and May month has the highest numbers of sales.

- There was a drastic decrease in the year 2020 of February month.

- We can not determine the reason behind it as the records shows only the data for two months.
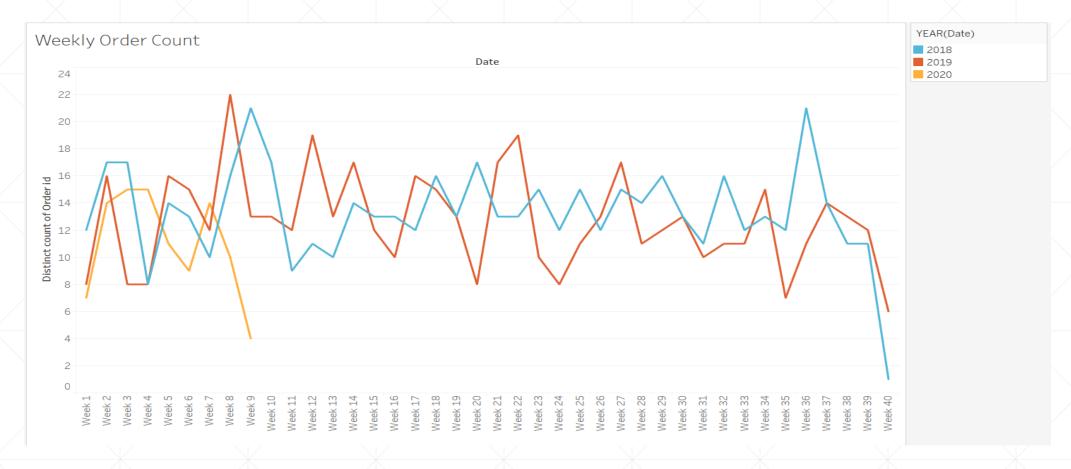
MONTHLY ORDERS EXCLUDING 2020

Monthly Order Count 2018 & 2019

- May month brings good orders followed by March and February.

- August 2019 had received the lowest order of 48.

- There is significant drop in orders between May - June and March – April.

WEEK-WISE ORDERS TREND

Weekly Order Count

- In contrast to 2018 and 2019, week 5, 8 and 9 has seen a fall in orders for 2020 whereas, week 7 was the opposite. Week 20, 23 saw a rise in orders in 2018 whereas it saw a significant drop in 2019. Week 40 shows a big dip in the orders. The above patterns indicate lack of consistency in order placing.

DAY-WISE ORDERS TREND

- The highest distinct count of ordered date is 17.
- No proper trend can be analyzed using daily wise orders.

INFERENCES:

- 📊 **Dataset Structure & Coverage**

- The dataset comprises **20,641 records** across **3 primary features**.

- Data has been captured for the years **2018, 2019**, and **early 2020**, though coverage is uneven:
  - For 2018 and 2019, data is available only up to **September (Q3)**.
  - The year 2020 contains entries only for **January and February**, making its representation limited.

- 📄 **Orders and Product Inventory Overview**

- A total of **1,139 customer orders** have been placed within the available timeframe.

- There are exactly **37 different product types** in the inventory.

- **Poultry** is the **most frequently purchased item**, while **hand soap** appears at the bottom of the sales chart.

- **Soda** and **cereal** trail just behind poultry, while **sandwich loaves** and **fruits** are among the less popular products.

- When comparing product demand, the variation in order quantities across items isn't especially wide — there's a **fairly balanced distribution** overall.

INFERENCES:

- 📉 **Year-on-Year Order Trend Analysis**

- Overall, the store experiences a **modest volume of annual orders**, with no substantial spikes.

- There's a **noticeable decline in total order volume** as the years progress — 2019 lags behind 2018, and 2020 includes data for only two months.

- While **Q3 of 2018** stood out for having the **highest number of orders**, the same quarter in **2019 saw a major drop**.

- 🗓️ **Monthly & Seasonal Dynamics**

- **January** ranks highest in terms of monthly orders — likely boosted by 2020's inclusion.

- **May consistently performs well**, with **67 orders in May 2018** and **66 in May 2019**.

- Conversely, **June and August** (particularly in 2019) show a clear dip in activity.

- Overall, **March and February** also yield good performance — indicating a **mild seasonal pattern**.

INFERENCES:

- 📅 **Weekly Insights**

- There's clear **volatility in weekly order patterns**, with no uniform trends.

- For example:
  - **Weeks 5, 8, and 9 of 2020** show a slump, but **Week 7** sees a surge.
  - **Weeks 17, 26, 30, and 37** had growth in 2019, yet dropped during the same weeks in 2018.
  - In contrast, **Weeks 20 and 23** were strong in 2018 but saw reduced activity the following year.
  - Such week-wise differences underscore a **lack of consistent consumer demand** — possibly influenced by external or seasonal factors.

📅 **Day-Wise Patterns**

On a daily level, **the 17th and 30th** of each month seem to attract the most customer activity.
Most days record between **10 to 20 orders**, suggesting **steady but moderate day-to-day performance**.

INFERENCES:

- 🧠 **Overall Inference**

- The dataset hints at **sporadic customer engagement** and **non-uniform demand trends**.

- There are **elements of seasonality** present, but irregularities in month-to-month and week-to-week patterns suggest **potential external factors** influencing purchases.

- Opportunities lie in investigating **why certain products dominate** and **what drives specific time-bound order surges or declines**.

MARKET BASKET ANALYSIS

MARKET BASKET ANALYSIS AND ADVANTAGES
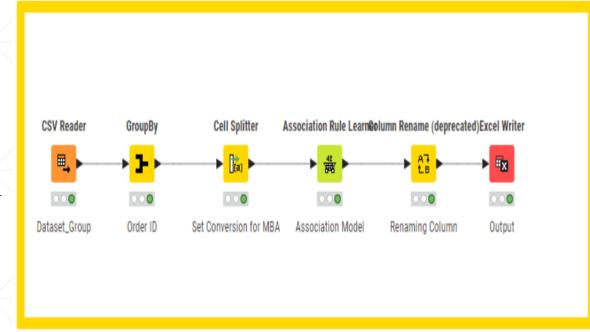
- Market Basket Analysis (MBA) is a data mining technique used by retailers and businesses to understand consumer buying patterns/behaviors to design strategies for increasing revenue and profits.

- Its main purpose is to identify the products which are purchased frequently in combination with other items and how strongly these combination of purchases can help increase the revenues. For eg; If consumer buys products A & B then he/she would buy C as well.

- Using Apriori Algorithm to do Market Basket Analysis of Customers purchasing behaviors. It can predict what the customer is going to buy next by looking at the products he is buying.

THRESHOLDS

- SUPPORT: This helps in identifying the frequency of occurrence for a set of products in an order or which set of products are purchased together the most out of the total purchases made. Set of products with higher support are likely to be purchased more frequently than the one's with a lower support.

- **For our project we have set minimum support as 0.05**

- CONFIDENCE: This denotes the probability of purchases containing recommended products (Consequent) in the same purchases already containing the set of products (Antecendent). For eg; How many purchases which contain product A (set of products) also contain product B (consequent). A higher confidence means its most likely that B will also be purchased if A is present in the order.

- **For our project we have set minimum confidence as 0.05**

- LIFT: This suggests the probability of purchasing recommended products if set of products are already present in order and if there is any association between the two. If lift is > 1, then the probability of purchasing recommended products increases if set of products are present in the order and it can be said that there is an association between the two and vice versa. A higher lift indicates higher level of association.

KNIME WORKFLOW AND OUTPUT TABLE

- Data is Grouped by Order_id

- Products are concatenated

- Set of Products are created using Cell Splitter

- Support is set at minimum 0.05

- Confidence is set at minimum 0.5

- We got a total of 1247 rules

- Consequent and Set Columns are renamed to Recommended_Product and Set_of_Products respectively

Rows: 1139 | Columns: 3

#	RowID	Order_id Number (integer)	Date String	Product String
7	Row6	7	03-01-2018	individual meals, paper towels, tortillas, milk, ice cream, juice, dishwashing
8	Row7	8	04-01-2018	ice cream, juice, paper towels, waffles, soda, cheeses, poultry, toilet paper
9	Row8	9	04-01-2018	juice, poultry, coffee/tea, coffee/tea, dishwashing liquid/detergent
10	Row9	10	05-01-2018	ketchup, coffee/tea, toilet paper, pork, flour, milk, soda, dishwashing liquid/(
11	Row10	11	05-01-2018	sandwich loaves, ice cream, soda, bagels, dishwashing liquid/detergent, eg(
12	Row11	12	06-01-2018	pork, tortillas, pork, shampoo, lunch meat, pasta, juice, bagels, bagels, launc
13	Row12	13	07-01-2018	sugar, fruits, all- purpose, aluminum foil, laundry detergent, individual meals
14	Row13	14	07-01-2018	fruits, dinner rolls, individual meals, shampoo, ketchup, cereals, sandwich b
15	Row14	15	07-01-2018	individual meals, ice cream, cereals, paper towels, bagels, mixes, lunch mea
16	Row15	16	08-01-2018	sugar, sandwich bags, flour, juice, milk, paper towels, cereals, sandwich bag
17	Row16	17	08-01-2018	milk, hand soap, pasta, individual meals, spaghetti sauce, cereals, sandwich
18	Row17	18	08-01-2018	sandwich bags, toilet paper, bagels, shampoo, coffee/tea
19	Row18	19	09-01-2018	individual meals, laundry detergent, coffee/tea, eggs, aluminum foil, beef, ju
20	Row19	20	10-01-2018	shampoo, dishwashing liquid/detergent, yogurt, juice, sugar, soap, sandwicl
21	Row20	21	11-01-2018	waffles, fruits, all- purpose, pork, juice, bagels, mixes
22	Row21	22	11-01-2018	cheeses, cereals, sugar, bagels, soda
23	Row22	23	11-01-2018	aluminum foil, bagels, shampoo, shampoo, dishwashing liquid/detergent, c(
24	Row23	24	11-01-2018	fruits, all- purpose, pasta, cheeses, juice, sandwich bags, sandwich loaves, (
25	Row24	25	11-01-2018	bagels, sugar, pork, sandwich loaves, tortillas, ice cream, all- purpose, yogur
26	Row25	26	12-01-2018	fruits, sandwich loaves, coffee/tea, aluminum foil, shampoo, cereals, dinner
27	Row26	27	13-01-2018	laundry detergent, pork, pasta, cheeses, fruits, sugar, lunch meat, laundry de
28	Row27	28	13-01-2018	pork, bagels, poultry, pasta, butter, all- purpose, pasta, shampoo, sugar, ketc
29	Row28	29	13-01-2018	pasta, butter, sandwich loaves, spaghetti sauce, juice, dinner rolls, all- purpc
30	Row29	30	14-01-2018	flour, bagels, cheeses, sandwich loaves, toilet paper
31	Row30	31	15-01-2018	aluminum foil, eggs, ice cream, pasta, juice, waffles, shampoo, dinner rolls,
32	Row31	32	15-01-2018	soap, paper towels, individual meals, dinner rolls, lunch meat, sugar, soap, e
33	Row32	33	15-01-2018	sandwich loaves, pork, sandwich bags, ketchup, coffee/tea, soda, poultry, p
34	Row33	34	16-01-2018	lunch meat, mixes, soap, hand soap, tortillas, coffee/tea, cheeses, tortillas, (
35	Row34	35	16-01-2018	aluminum foil, cheeses, cereals, mixes, laundry detergent, juice, pork, bagel:

ASSOCIATION RULES

- Association Rule determines the frequency of occurrence of an itemset in a purchase order where itemset represents a set of products.

- It helps show how strong or weak the associations between the products are. It determines that products having strong associations are often purchased together in most cases and products with weak associations are said to be independent of each another.

- Using the association rules we would be able to determine which products have strong associations and which are often purchased together helping our store owner to form cross-selling strategies, promotional or/and combo or/and discount offers, shuffle with in-store product placement etc.

- Moreover, the association rules are proven to be very useful for retailers which again makes it very relevant for our grocery store to make use of the same and increase revenue and profits.

ASSOCIATION RULES

- When you apply Association Rule Mining on a given set of transactions T your goal will be to find all rules with:

1. Support greater than or equal to min_support

2. Confidence greater than or equal to min_confidence

- There are 20,642 records in the grocery dataset, totalling 1139 orders.

- Customers have purchased numerous goods in a single order, with an average of 18 products per buy.

- As a result, it is critical to investigate the association between two items.

- Once the relationship has been assessed, actions can be taken based on the confidence and lift indicators to boost sales by offering combos and giftpacks.

ASSOCIATION RULE OUTPUT

Renamed/Retyped table (Table) — □ ▢ ✕

Rows: 10550 | Columns: 6

#	RowID	Support Number (double)	Confidence Number (double)	Lift Number (double)	Suggested Item String	implies String	Basket Items String
1	rule0	0.031	0.574	1.469	soda	<---	[yogurt, toilet paper, pork]
2	rule1	0.031	0.565	1.468	yogurt	<---	[flour, juice, sandwich bags]
3	rule2	0.031	0.556	1.445	yogurt	<---	[fruits, flour, sandwich bags]
4	rule3	0.031	0.593	1.543	yogurt	<---	[flour, sandwich bags, sugar]
5	rule4	0.031	0.565	1.468	yogurt	<---	[dishwashing liquid/detergent, juice,
6	rule5	0.031	0.565	1.468	yogurt	<---	[tortillas, sandwich bags, sugar]
7	rule6	0.031	0.614	1.554	lunch meat	<---	[yogurt, ice cream, hand soap]
8	rule7	0.031	0.556	1.406	lunch meat	<---	[yogurt, fruits, tortillas]
9	rule8	0.031	0.556	1.445	yogurt	<---	[dinner rolls, all- purpose, toilet paper
10	rule9	0.031	0.556	1.445	yogurt	<---	[all- purpose, milk, sugar]
11	rule10	0.031	0.565	1.468	yogurt	<---	[shampoo, flour, beef]
12	rule11	0.031	0.574	1.492	yogurt	<---	[flour, tortillas, beef]
13	rule12	0.031	0.574	1.492	yogurt	<---	[sandwich loaves, flour, hand soap]
14	rule13	0.031	0.574	1.492	yogurt	<---	[spaghetti sauce, flour, sugar]
15	rule14	0.031	0.565	1.451	dinner rolls	<---	[yogurt, butter, spaghetti sauce]
16	rule15	0.031	0.565	1.502	mixes	<---	[yogurt, butter, spaghetti sauce]
17	rule16	0.031	0.556	1.445	aluminum foil	<---	[yogurt, fruits, hand soap]
18	rule17	0.031	0.556	1.445	yogurt	<---	[shampoo, spaghetti sauce, milk]
19	rule18	0.031	0.556	1.445	yogurt	<---	[fruits, mixes, juice]
20	rule19	0.031	0.565	1.616	sandwich loaves	<---	[yogurt, hand soap, bagels]
21	rule20	0.031	0.556	1.59	sandwich loaves	<---	[yogurt, fruits, hand soap]
22	rule21	0.031	0.565	1.448	eggs	<---	[yogurt, hand soap, bagels]
23	rule22	0.031	0.565	1.52	pasta	<---	[paper towels, yogurt, hand soap]
24	rule23	0.031	0.574	1.492	yogurt	<---	[paper towels, cheeses, tortillas]
25	rule24	0.031	0.556	1.445	yogurt	<---	[dishwashing liquid/detergent, eggs,
26	rule25	0.031	0.574	1.492	yogurt	<---	[sandwich loaves, fruits, sugar]
27	rule26	0.031	0.603	1.64	sandwich bags	<---	[laundry detergent, beef, pork]
28	rule27	0.031	0.556	1.482	beef	<---	[sandwich loaves, pork, sandwich ba

Threshold Values of lift values, Support and Confidence

- Support and Confidence measure how interesting the rule is. It is set by the minimum support and minimum confidence thresholds.

- These thresholds set by client help to compare the rule strength according to your own or client's will.

- The closer to threshold the more the rule is of use to the client.

- **Frequent Item sets:**

  - Item-sets whose support is greater or equal than minimum support threshold (min_sup).

- **Strong rules:**

  - If a rule A=>B[Support, Confidence] satisfies min_sup and min_confidence then it is a strong rule.

- **Lift:**

  - Lift gives the correlation between A and B in the rule A=>B.

  - Correlation shows how one item-set A effects the item-set B.

  - Lift(A=>B)=SupportSupp(A)Supp(B)

- A rule may appear to have a strong association in a data collection because it appears frequently, but it may emerge much less frequently when applied.

Lift values, Support and Confidence

- There is no particular threshold value of support. If the dataset is large then it is advisable to set the value at 10% and increase it accordingly till required number of associations are generated.

- The threshold value of support in this case is 10%. Association rules are not developing if the value goes beyond 10%

- Also, the confidence is kept at 40% which is optimal level.

- The associations with highest lift values are considered to be accepted more.

- In the following case, the association rule is formed highest lift value that is 12.34% at a confidence level of 46% and support of 18%.

- This rule implies that the customers who bought dishwashing also purchased mixes.

- Also the customers who bought soda have high chances of buying eggs.

INTERPRETATION OF ASSOCIATION RULES OUTPUT

Row ID	D Support	D Confide...	D ▼ Lift	S Recommended_Product	S implies	S Set_Of_Products
rule59	0.055	0.649	1.791	paper towels	<---	[eggs, ice cream, pasta]
rule58	0.055	0.643	1.731	pasta	<---	[paper towels, eggs, ice cream]
rule21	0.051	0.674	1.726	cheeses	<---	[bagels, cereals, sandwich bags]

- In Rule 59, there is support that 5.5% orders have eggs, ice cream, pasta and paper towels. If eggs, ice cream and pasta are purchased in an order, then there is confidence of 64% that paper towels will also be purchased in the same order or and there is 79% chance that wherever eggs, ice cream and pasta are in an order, paper towels will also be purchased.

- In Rule 58, there is support that 5.5% orders have eggs, ice cream, pasta and paper towels. If eggs, ice cream and paper towels are purchased in an order, then there is confidence of 64% that pasta will also be purchased in the same order and there is 73% chance that wherever eggs, ice cream and paper towels are in an order, pasta will also be purchased.

- In Rule 21, there is support that 5.5% orders have bagels, cereals, sandwich bags and cheeses. If bagels, cereals, sandwich bags are purchased in an order, then there is confidence of 67 % that cheeses will also be purchased in the same order and there is 73% chance that wherever bagels, cereals, sandwich bags are in an order, cheeses will also be purchased.

RECOMMENDATIONS – KNIME ANALYSIS

- From the analysis, one can conclude that the number of orders has been falling drastically over the years.

- The highest number of orders was in 2018 and then followed by 2019.

- The Q4 data has not been provided or can conclude that there are no orders placed in Q4.

- The store can provide discounts in the middle of the month in order to attract more customers.

- Poultry and ice-creams are most frequently ordered items. It is advisable to increase the variety of items in this category so that customers will have a lot to choose.

- The months of mid quarters have been showing a consistent performance over the years.

- Also one can conclude that the numbers of orders are high in starting and end days of the months.

- This can be because most of the customers get their pay either at the starting or at the end of every month.

- Hand soaps and loaves are least purchased items. So the store can invest a bit less on these items.

- In order to increase the sales, Q4 is crucial part for many businesses as it is festive season customers tend to order a lot.

- It is advisable to provide the service in Q4 as well.

- The decrease in the orders over the years can be of many reasons like poor customer service or no proper offers provided. It is recommended to look into the service and provide customers the highest satisfaction in terms of service as well as products.

RECOMMENDATIONS – KNIME ANALYSIS

Row ID	D Support	D Confide...	D ▼ Lift	S Recommended_Product	S implies	S Set_Of_Products
rule59	0.055	0.649	1.791	paper towels	<---	[eggs, ice cream, pasta]
rule58	0.055	0.643	1.731	pasta	<---	[paper towels, eggs, ice cream]

- Can roll out Combo offers for eggs, ice cream, pasta and paper towels or can offer 5% discount on Pasta if paper towels, eggs and ice cream are already purchased as pasta has a lower rate of order.

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule21	0.050921861	0.674418605	1.726208518	cheeses	<---	[bagels, cereals, sandwich bags]
rule20	0.050921861	0.610526316	1.659640749	sandwich bags	<---	[cheeses, bagels, cereals]

- Can roll out Combo offers for bagels, cereals, sandwich bags and cheeses or can offer 15% discount on sandwich bags if bagels, cereals and cheeses are already purchased as sandwich bags have a lower rate of order.

RECOMMENDATIONS – KNIME ANALYSIS

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule18	0.050921861	0.630434783	1.677722471	mixes	<---	[yogurt, poultry, aluminum foil]

- Can roll out Combo offers for yogurt, poultry, aluminium foil and mixes

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule40	0.051799824	0.641304348	1.648861517	dinner rolls	<---	[spaghetti sauce, poultry, ice cream]

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule24	0.050921861	0.630434783	1.620914712	dinner rolls	<---	[spaghetti sauce, poultry, cereals]

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule52	0.053555751	0.642105263	1.650920756	dinner rolls	<---	[spaghetti sauce, poultry, laundry detergent]

- Can roll out Combo offers for dinner rolls, spaghetti sauce, poultry and ice cream/cereals/laundry detergent or can also provide free dinner rolls if spaghetti sauce, poultry and ice cream/cereals/laundry detergent are purchased. This can help improve the sale of spaghetti sauce as the other items are purchased regularly and spaghetti sauce would cover the cost of dinner rolls.

RECOMMENDATIONS – KNIME ANALYSIS

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule11	0.050043898	0.612903226	1.615964755	coffee/tea	<---	[yogurt, cheeses, cereals]

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule8	0.050043898	0.587628866	1.528103375	yogurt	<---	[cheeses, cereals, coffee/tea]

- Can roll out Breakfast Combo for coffee/tea, yogurt, cheeses and cereals during the weeks and months with lower sales. Also these products can be placed close to each other to increase the chance of being purchased together. Can offer 10% discount on Yogurt if cheeses, cereals and coffee/tea is purchased.

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule870	0.082528534	0.562874251	1.49929375	individual meals	<---	[sandwich loaves, lunch meat]

- Can roll out Lunch Meal Combo offers for individual meals, sandwich loaves and lunch meat.

RECOMMENDATIONS – KNIME ANALYSIS

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule49	0.051799824	0.634408602	1.627458103	eggs	<---	[paper towels, dinner rolls, pasta]

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule35	0.051799824	0.627659574	1.610144719	eggs	<---	[dinner rolls, poultry, soda]

- Can roll out Combo offers for eggs, paper towels, dinner rolls and pasta as seen in rule49 or eggs, dinner rolls, towels, poultry and soda as seen in rule35. Can also offer 10% discount on eggs if purchased with either of the two rules.

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule28	0.050921861	0.574257426	1.517585169	laundry detergent	<---	[poultry, milk, cereals]

- Can offer 10% discount on laundry detergent if poultry, milk and cereals are already purchased. This way, we can try and improve sales of laundry detergent using the combo of the other products. We can also offer all four products at a reasonable price.

RECOMMENDATIONS – KNIME ANALYSIS

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule909	0.082528534	0.513661202	1.470000275	sandwich loaves	<---	[cheeses, ketchup]

▪ Can roll out Combo offers for sandwich loaves, cheeses and ketchup while also placing them near each other.

Row ID	Support	Confidence	Lift	Recommended_Product	implies	Set_Of_Products
rule32	0.051799824	0.567307692	1.452052723	soda	<---	[eggs, dinner rolls, poultry]

▪ Can roll out Combo offers for soda, eggs, dinner rolls and poultry.

OFFERS:

- Also most of the customers are inclined to buy ketchups along with sandwich bags, the store can provide an offer saying get 1 ketch up free with 2 sandwich bags

- An offer like buy 1 kg of beef and get 1 kg of assorted fruits free can be provided to the customers.

- Since hand soaps are least preferred by customers they can be combined with any of the highest selling product.

- Hand soaps can be combined with detergents to increase the purchase of hand soaps as well.

- Sandwich loaves if also one of the least selling products. It is evident that people who buy loaves also buy individual meals. So the store can provide combo offer like buy three meals and get 5 sandwich loaves free.

- In order the increase the sales, the grocery store can provide combo offers to its customers.

- **Top 5 combos according to Market Basket Analysis are:**

1. **Mixed & Dishwashing liquid**
2. **Eggs & Soda**
3. **Juice & Shampoo**
4. **Juice and Spaghetti sauce**
5. **Pasta & Paper towels**

THANK YOU
HIYA SHAH
hiyashah2020@gmail.com