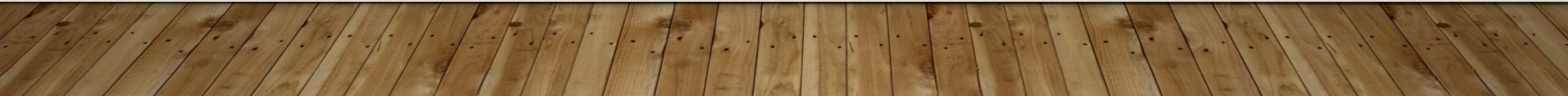# MRA PROJECT-2

ISSAC ABRAHAM

CONTENTS

- Exploratory Analysis
- Market Basket Analysis
- Associations Identified
- Suggestion of Combos

EXECUTIVE SUMMARY

- The grocery store transactional dataset offers valuable insights into customer purchase behavior and presents opportunities for revenue growth through strategic combo offers and discounts. By analyzing the Point of Sale (POS) data, we identified patterns in the most frequently occurring item sets in customer orders. Leveraging market basket analysis techniques, we derived association rules to suggest lucrative combo offers tailored to customer preferences. Our recommendations aim to enhance customer satisfaction, drive sales, and optimize revenue for the grocery store.

SHAPE ,INFO & SUMMARY OF THE DATASET

- The dataset contains 20,641 entries and 3 columns. It includes information about the date of orders, order IDs, and the products purchased. The Date column is of type object, while Order_id is of type integer, and Product is of type object.

- The order ID column has a total count of 20,641 entries, with a mean value of approximately 576 and a standard deviation of about 329. The minimum order ID is 1, while the maximum is 1139, with the median value at 581.

	Order_id
count	20641.000000
mean	575.986289
std	328.557078
min	1.000000
25%	292.000000
50%	581.000000
75%	862.000000
max	1139.000000

```
<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
```

```
(20641, 3)
```

YEARLY & MONTHLY TRENDS IN SALES

- There was maximum sales in the year 2018 followed by the year 2019 and a steep decline in the year 2020 by more than 60% from the previous years.

- There is highest sales in the first month leading to a u shaped pattern in sales during the other months. The sales increase gradually after month of may.

Yearly Trend

Monthly Trend
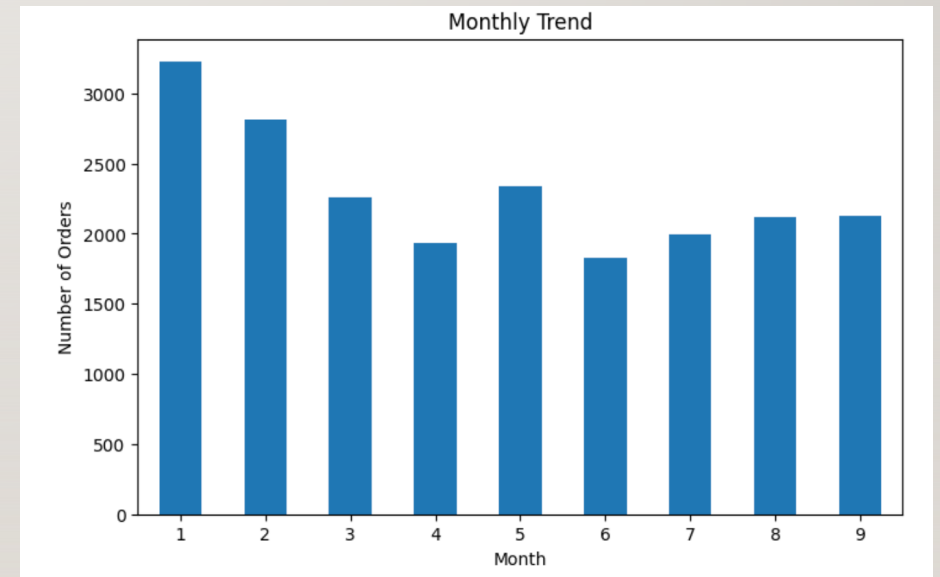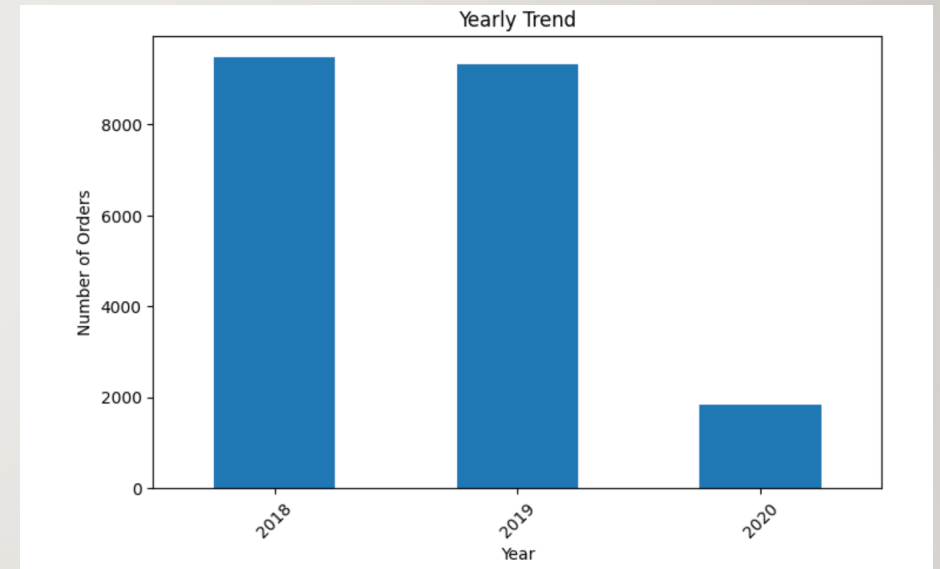
QUARTERLY AND DAILY SALES TREND

- There is highest sales in the first Quarter and the second highest sales in the 3rd quarter leading to a decline of 20 percent.

- We can see there is highest sales on the weekends and the lowest sales on Monday.


Quarterly Trend

Daily Trend
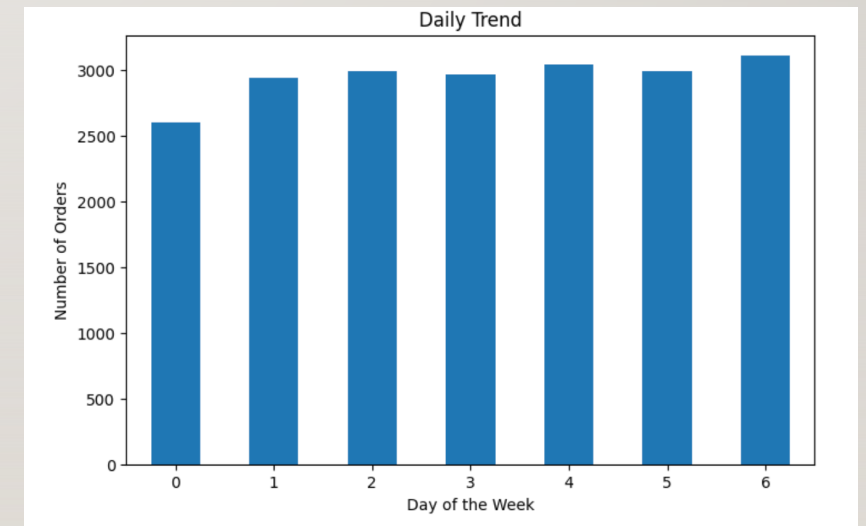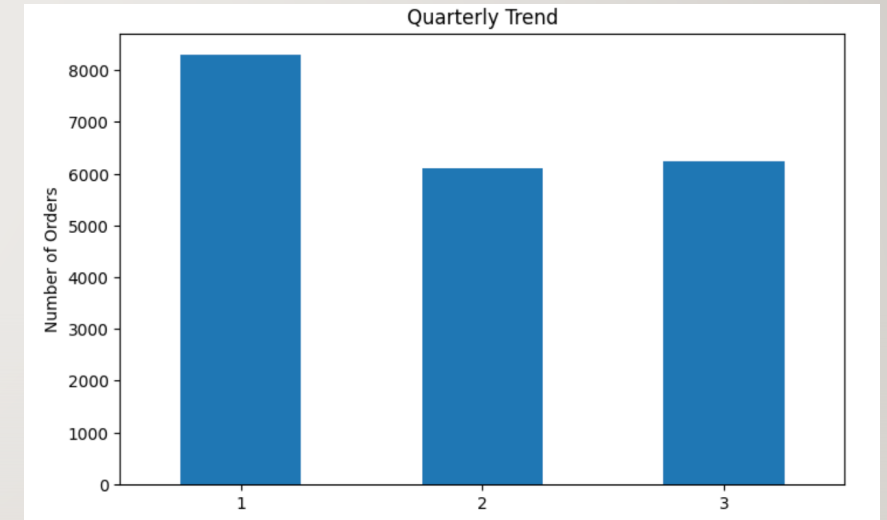
MARKET BASKET ANALYSIS (ASSOCIATION)

- Association rules play a crucial role in market basket analysis, revealing underlying connections and trends among items commonly bought simultaneously.

- In the context of a grocery store, these rules offer insights into customer shopping habits, aiding in strategic product arrangement, tailored promotional strategies, and enticing package deals.

- By harnessing association rules, retailers can elevate the shopping experience, enhance customer engagement, and ultimately increase sales revenue through encouraging supplementary purchases and maximizing cross-selling potential.

KNIME WORKFLOW

- Read CSV and group by order_id.

- Concatenated items, removed duplicates, and split the cells using the tool Knime.

- We set the minimum support threshold at 0.1, meaning that we only considered itemsets with a frequency of at least 10% in the dataset.

- The confidence threshold was set to 0.4, ensuring that only associations with a probability of occurrence of at least 40% were identified.
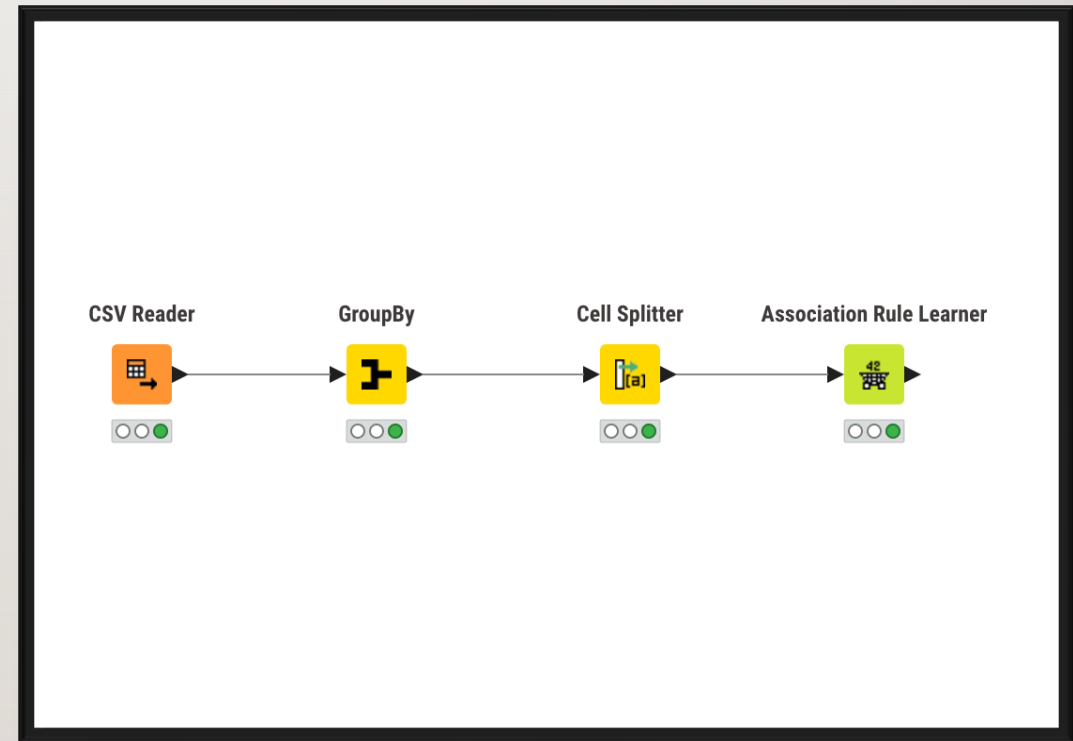
OUTPUT TABLE HEAD

- Support: Indicates the proportion of transactions that contain both the antecedent and consequent.

- Confidence: Denotes the probability of the consequent occurring given that the antecedent has occurred.

- Lift: Measures the strength of association between the antecedent and consequent. A lift value greater than 1 indicates a positive association.

Rule	Support	Confidence	Lift	Antecedent	Consequent
rule893	0.18	0.464	1.234	mixes	dishwashing liquid/detergent
rule909	0.187	0.48	1.228	soda	eggs
rule787	0.17	0.452	1.226	shampoo	juice
rule825	0.172	0.457	1.224	spaghetti sauce	juice
rule648	0.164	0.453	1.219	pasta	paper towels
rule871	0.176	0.469	1.218	yogurt	juice
rule493	0.16	0.457	1.217	individual meals	sandwich loaves
rule821	0.172	0.457	1.207	toilet paper	juice
rule630	0.164	0.442	1.202	sandwich bags	ketchup
rule879	0.177	0.461	1.199	yogurt	aluminum foil
rule831	0.173	0.466	1.197	dinner rolls	pasta
rule680	0.166	0.448	1.195	beef	fruits
rule881	0.177	0.462	1.195	lunch meat	individual meals

Table 1: Top associations

SUGGESTION OF COMBOS

Recommendation for possible combos with lucrative offers

Combo Offer	Combo Items	Offer
1	Mixes and Dishwashing Liquid/Detergent	Buy Mixes, Get 20% off Dishwashing Liquid/Detergent
2	Shampoo and Juice	Buy Shampoo, Get 15% off Juice
3	Spaghetti Sauce and Juice	Buy Spaghetti Sauce, Get 10% off Juice
4	Yogurt and Juice	Buy Yogurt, Get 20% off Juice
5	Pasta and Paper Towels	Buy Pasta, Get 15% off Paper Towels
6	Individual Meals and Sandwich Loaves	Buy Individual Meals, Get 15% off Sandwich Loaves
7	Toilet Paper and Juice	Buy Toilet Paper, Get 10% off Juice
8	Sandwich Bags and Ketchup	Buy Sandwich Bags, Get 10% off Ketchup
9	Yogurt and Aluminum Foil	Buy Yogurt, Get 15% off Aluminum Foil
10	Dinner Rolls and Pasta	Buy Dinner Rolls, Get 20% off Pasta

These offers are designed to encourage customers to purchase specific combinations of products by providing them with attractive discounts.