
Data Scientist

Marketing Analysis

Market Basket.

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been



● Problem.

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.

01 Overview

We provide a lay of the land view and describe the data we are working with

02 EDA

We run exploratory analysis and look at high level insights

03 RFM

We segment our customers using RFM and showcase our result

04 Recommendations

We run inferences and showcase our analysis and give recommendations to business

Lay of the land.

In this section we describe the high-level view of the data and where the business sit currently.



Missing Data

There is no missing data
in the analysis



**2yrs
11mths**

The data covers a
period of 2 years and 11
months



20 641

The data has 20 641
rows and 3 columns



37

There are 37 unique
products in the data



1139

There are 1139 unique
orders in the data



1.07

The store averages 1.07
orders per day

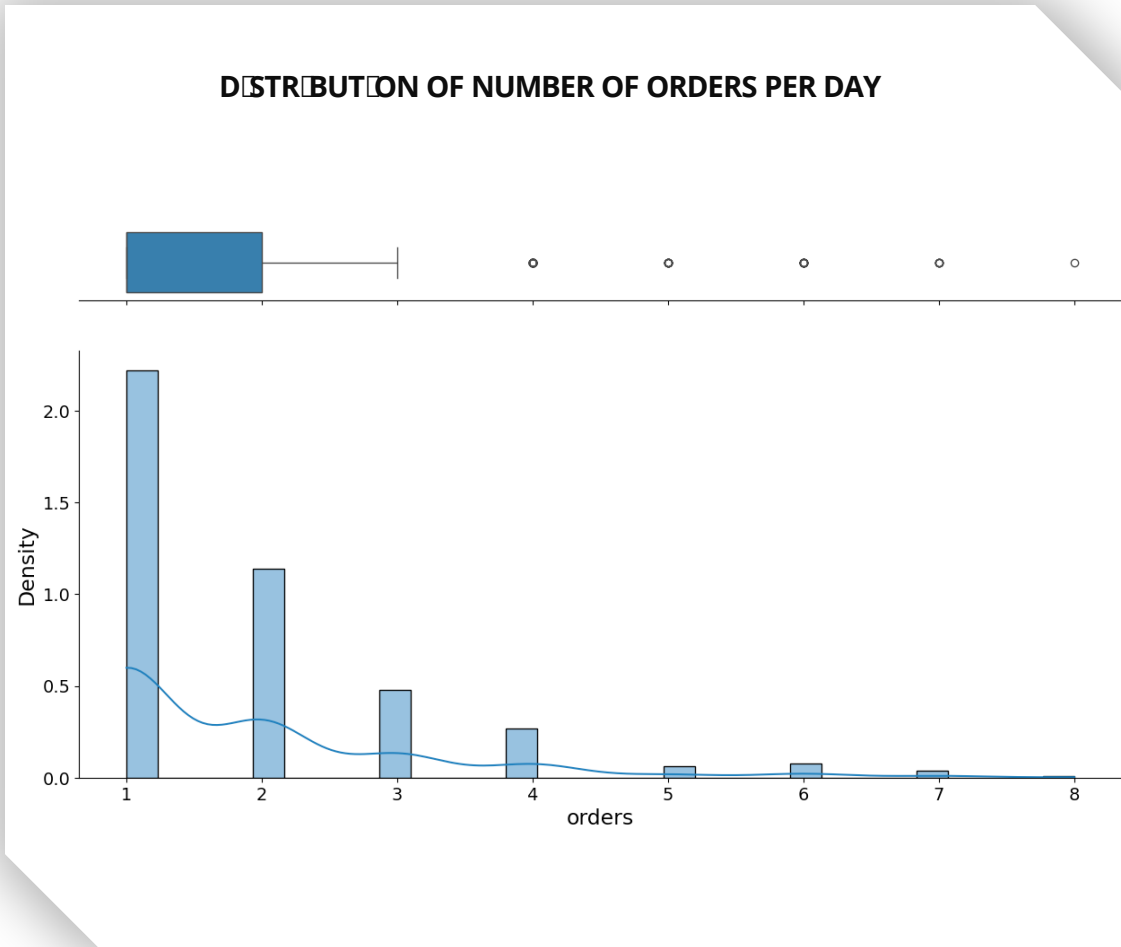
EDA

In this section we explore the data and find high level insights that speak to our data. We will undertake the following:

- 1, Uni-variate analysis
2. Mult-variate Analysis
3. Time Series Analysis

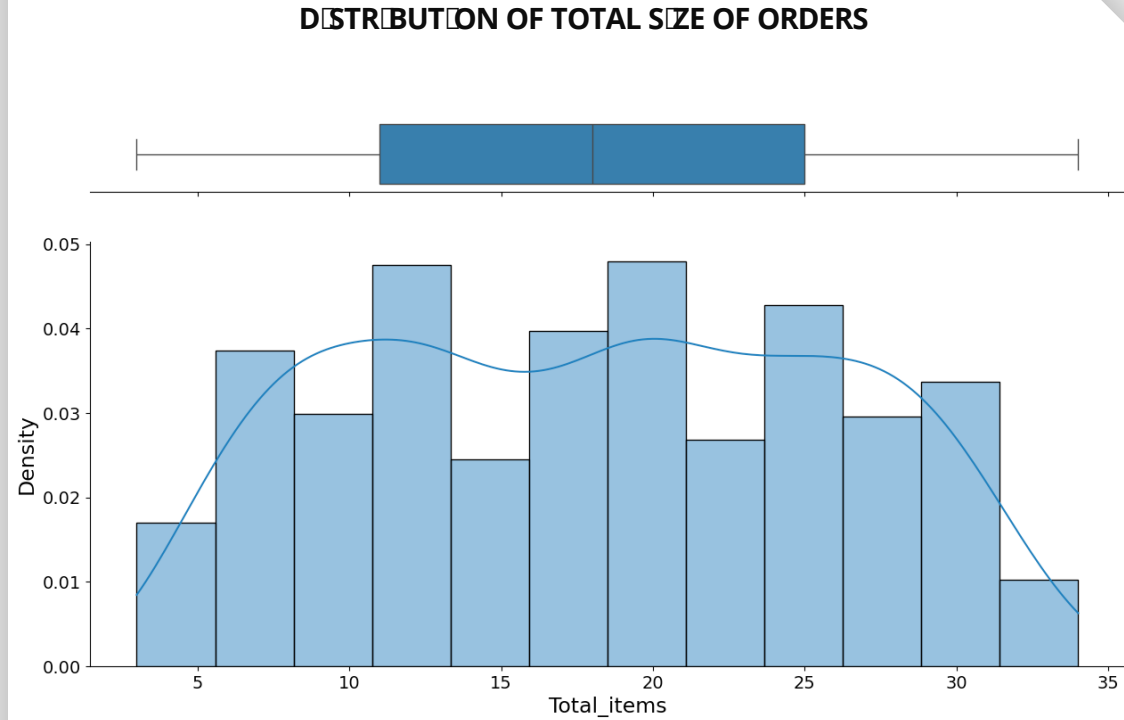
NUMBER OF ORDERS

We can see that our shop activity by day is quite low with most of the days having a single order. The store needs to run promotions to drive foot traffic and order conversion.



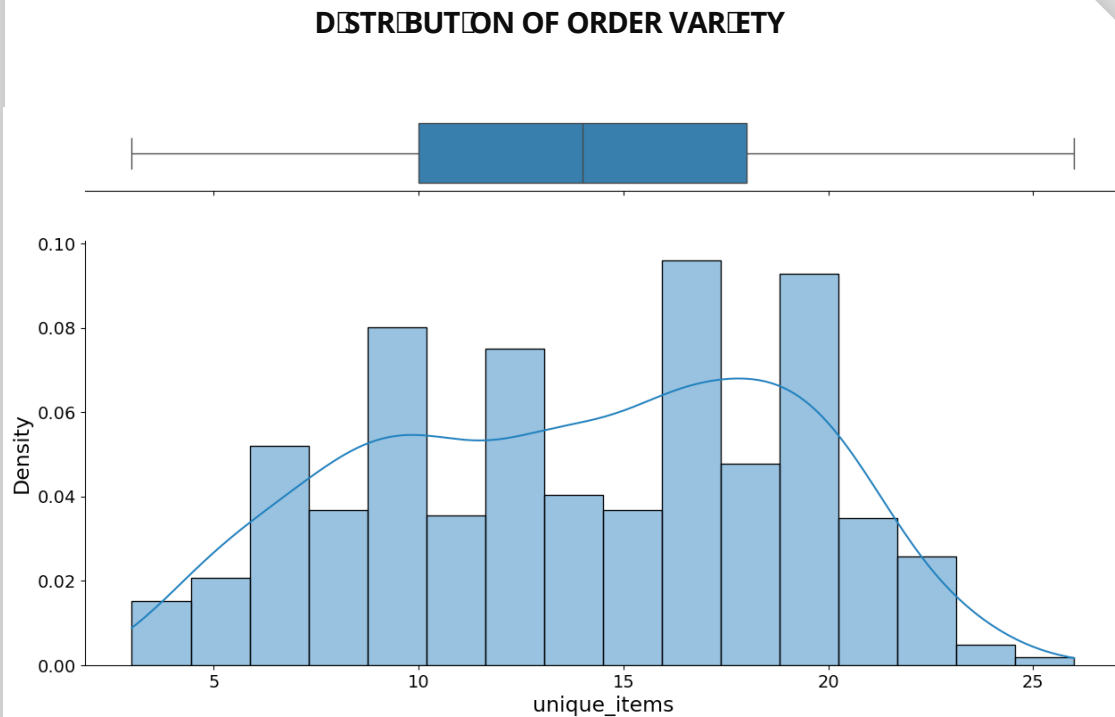
ORDER SIZE

Our order sizes tend to be quite large with most of the orders containing 10 to 25 items



ORDER VARIETY

Our order baskets tend to be quite varied with 10 to 18 unique items occurring in most orders



BUNDLE PROMOTIONS

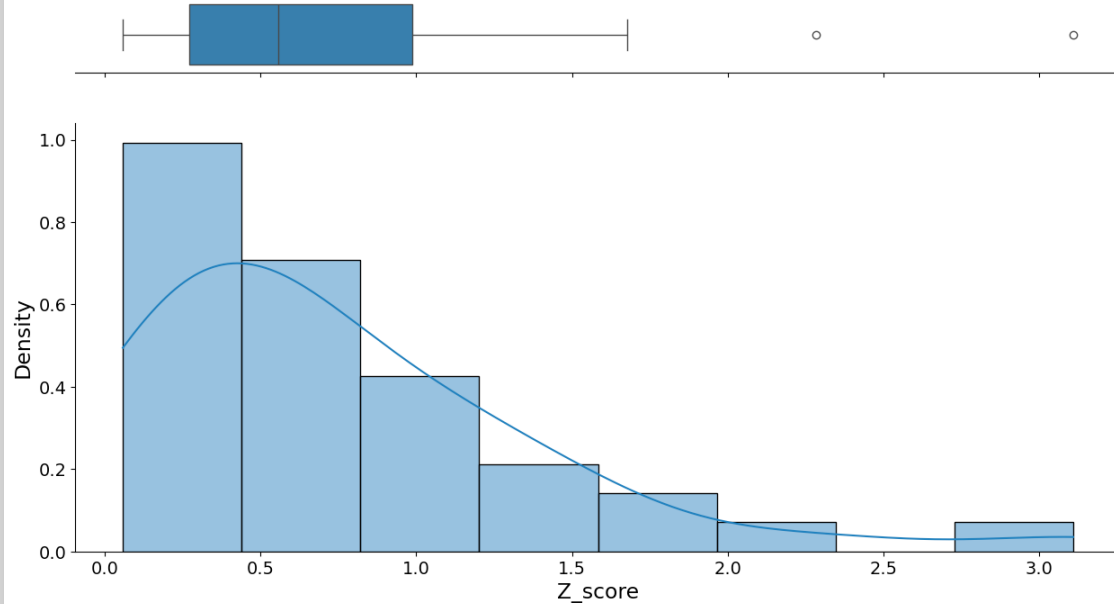
Depending on how we want to run promotions we have two options for the buy two get one free promotion. the analysis reveals two types of promotions:

1. An Aggressive promotion
2. A more conservative promotion strategy

- For the aggressive promotion we take products that have a Z-score greater than 1 (1 standard deviation from the mean)
these products will be (paper towels, toilet paper, individual meals, milk, spaghetti sauce, fruits, waffles, ice cream, all-purpose)

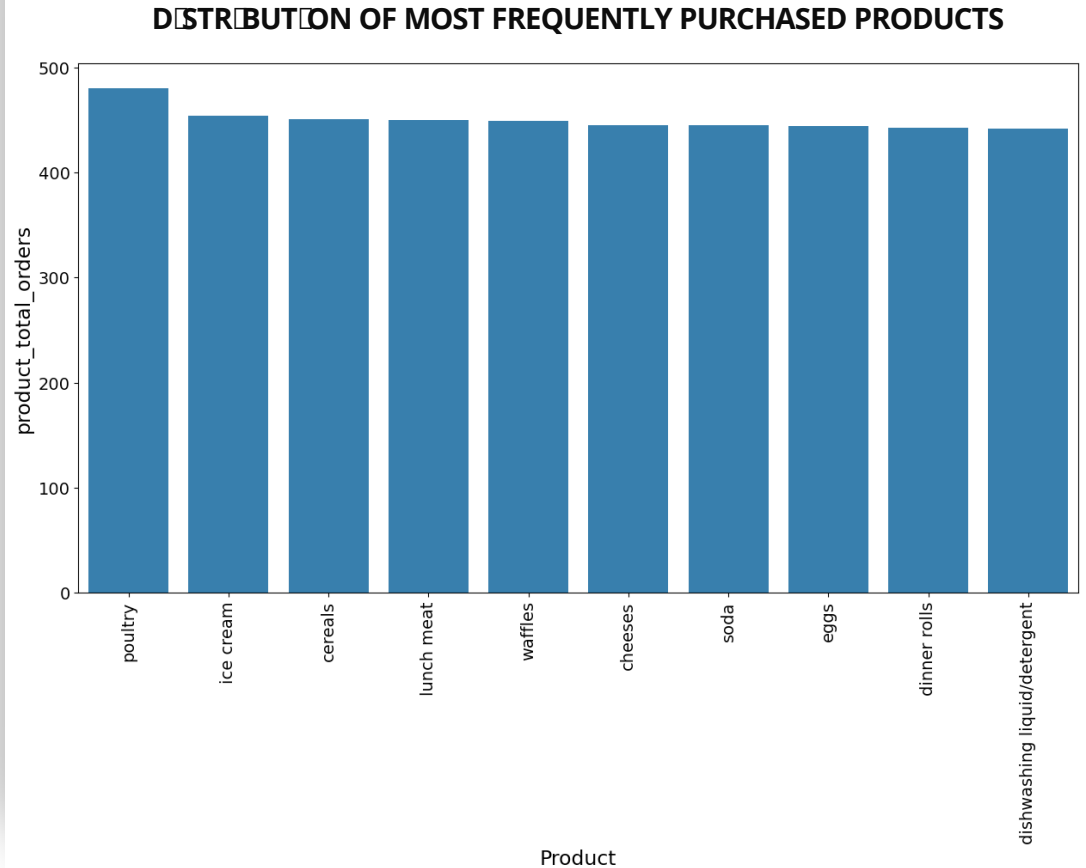
- For the aggressive promotion we take products that have a Z-score greater than 1.5 (1.5 standard deviation from the mean)
these products will be (paper towels, toilet paper, individual meals, milk)

DISTRIBUTION OF BUNDLED PRODUCT



MOST PURCHASED PRODUCTS

We can see that poultry is our most purchased product

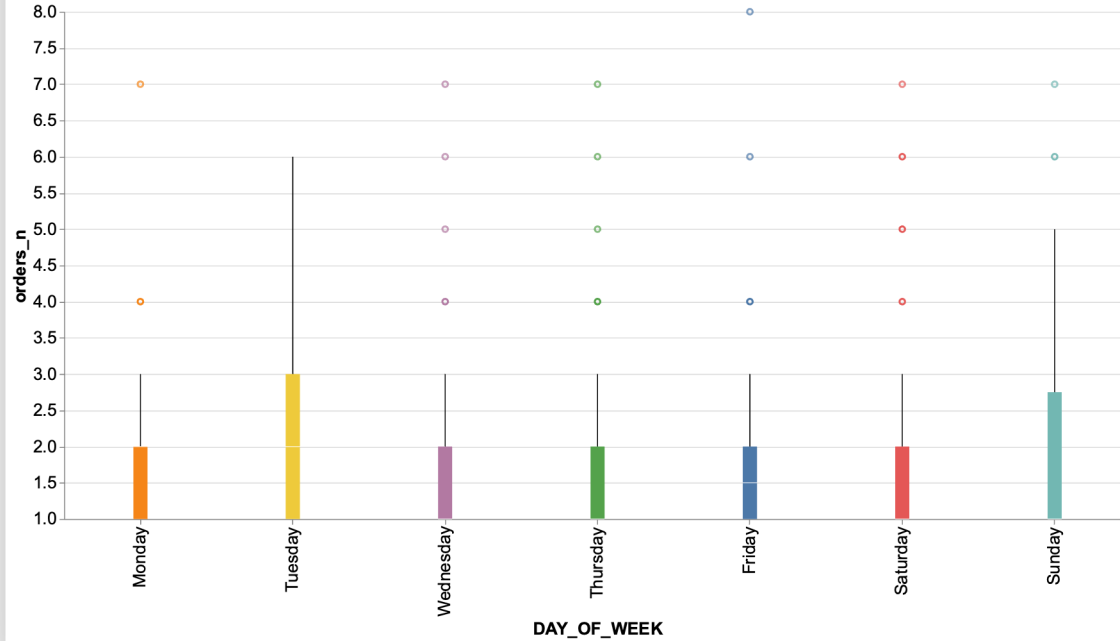


HIGHEST SALES DAY

We experience more orders on Tuesday and Sunday as we can see that the median orders for Tuesday are much higher than any other day.

The store is generally not busy.

DISTRIBUTION OF DAY WITH MOST ORDERS



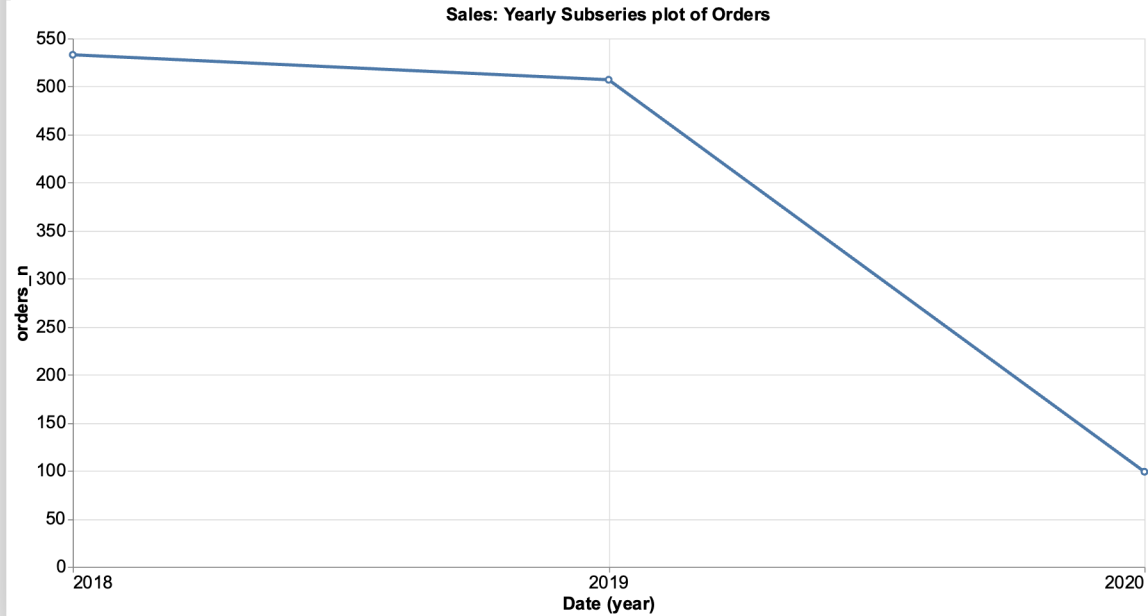


Data Scientist

How are we trending over time

SALES BY YEAR

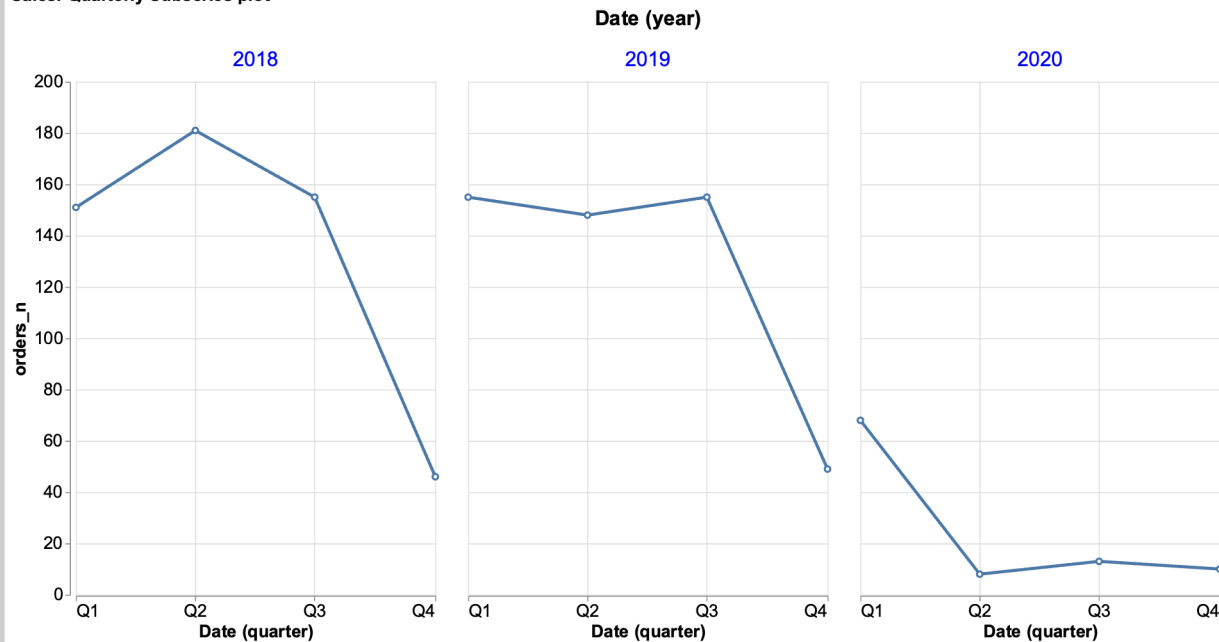
We can see that 2018 was our highest sales year



QUARTERLY TREND

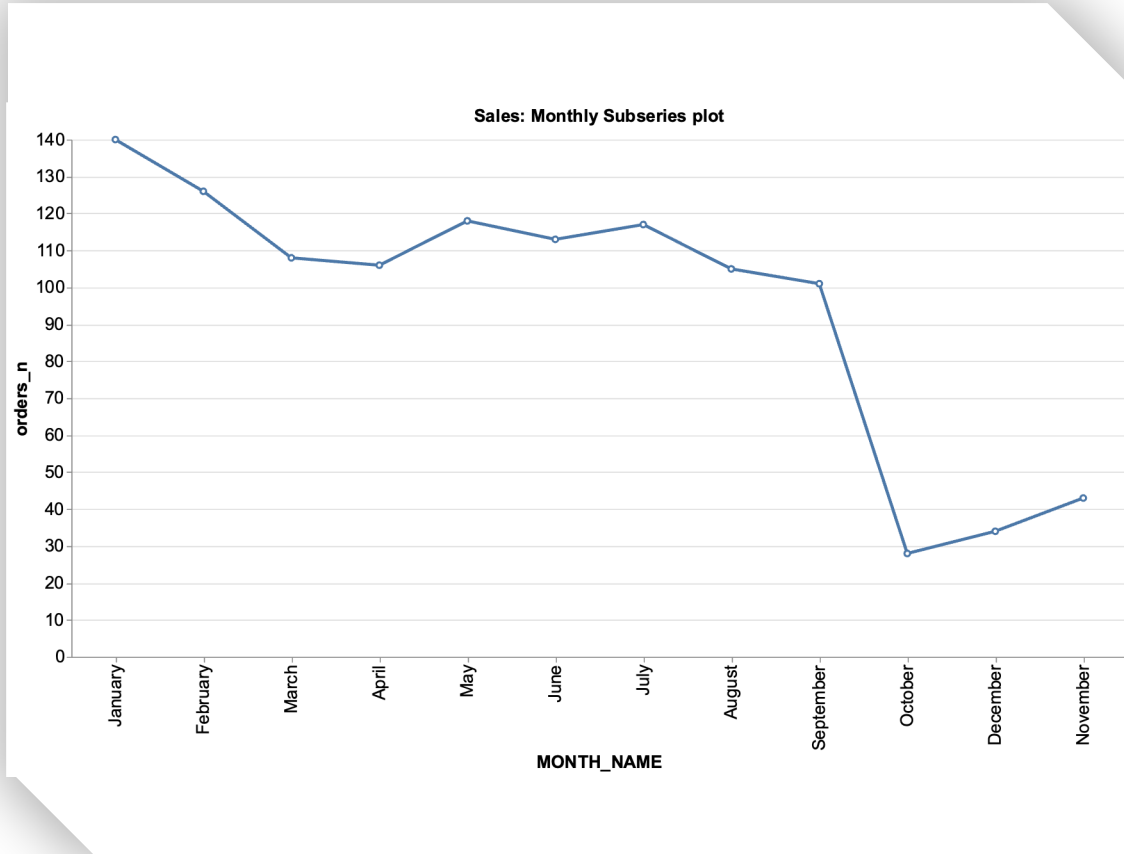
We can see a clear quarterly trend with Quarter 2 having the highest sales while Q4 has the lowest sales in the data

Sales: Quarterly Subseries plot



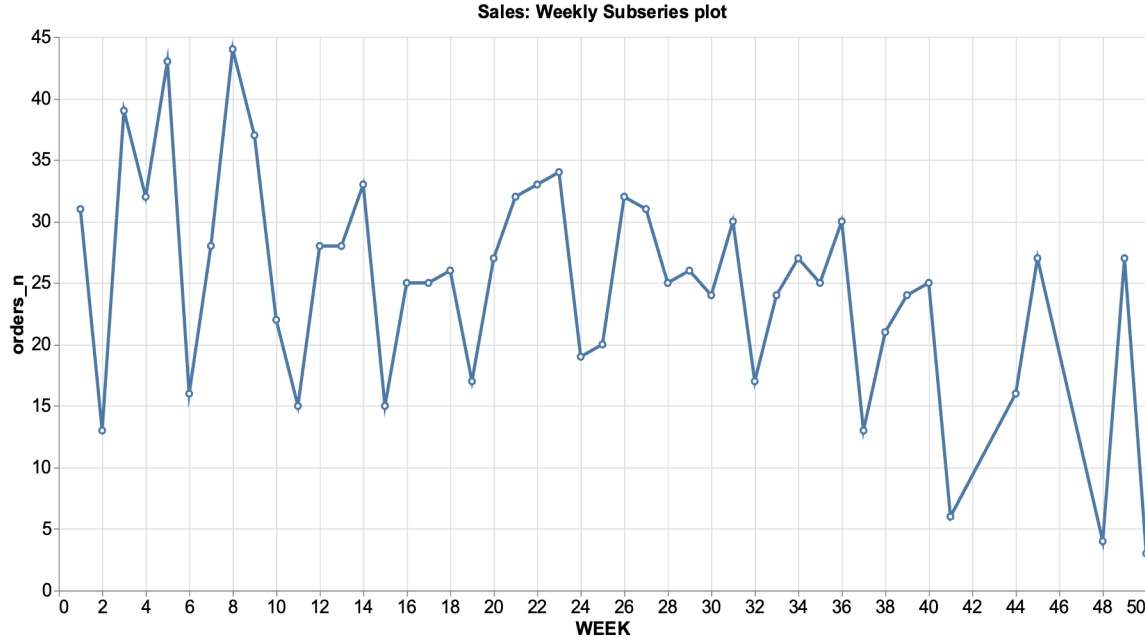
MONTHLY TREND

The monthly plot again speaks to the lower sales in Quarter 4 with a marked drop from September to October probably speaking to a holiday season where people travel elsewhere or some seasonal factor



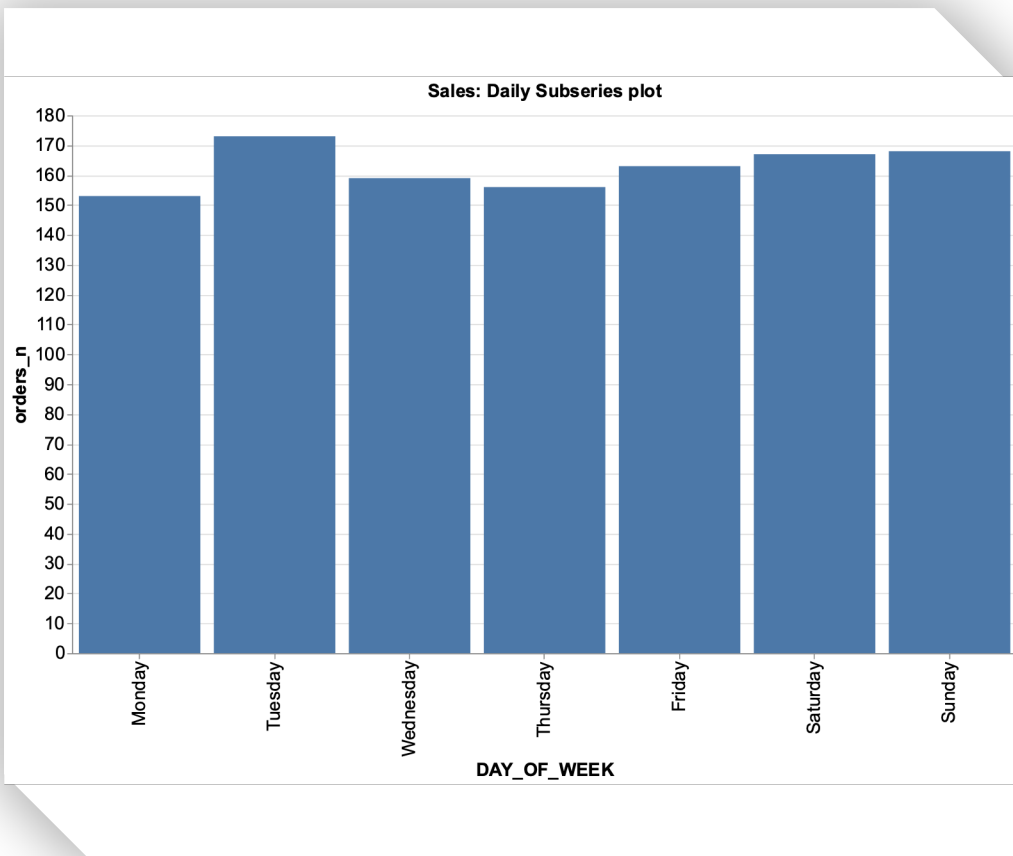
WEEKLY TREND

The weekly trend again speaks to the lower sales in quarter four and you can see a weekly pattern which speaks to clients buying patterns where they may buy groceries for a 1–2-week period



DAY OF WEEK ORDERS

The average number of orders across days of the week is generally the same with Tuesday having the highest and it seems to be our busiest day. We can probably run more in-store promotions on Tuesday.



Market Basket Analysis

Market Basket analysis Market basket analysis is based on discovering associations between items by using **association rules** which take the form of if-then relationships. To build an association rule, we should have at least one **antecedent** and one **consequent**..

Association rule mining is a technique used to identify patterns in large data sets. It involves finding relationships between variables in the data and using those relationships to make predictions or decisions. .

Parameters and Assumptions:

- **Support:** Indicates how often an itemset appears in the dataset. A higher support means the rule is based on more transactions, which can be important for ensuring the rule is significant and not just a fluke.
- **Confidence:** Measures the likelihood that the consequent item is purchased when the antecedent item is purchased. High confidence means that there's a strong predictive power or likelihood of the consequent item being bought given the antecedent item.
- **Lift:** Compares the observed frequency of the antecedent and consequent appearing together to the frequency expected if they were independent. A lift greater than 1 means that the antecedent and consequent appear together more often than expected if they were statistically independent, indicating a positive association.

We are required to produce four segments for our clients.

KNIME WORKFLOW

PARAMETERS USED:

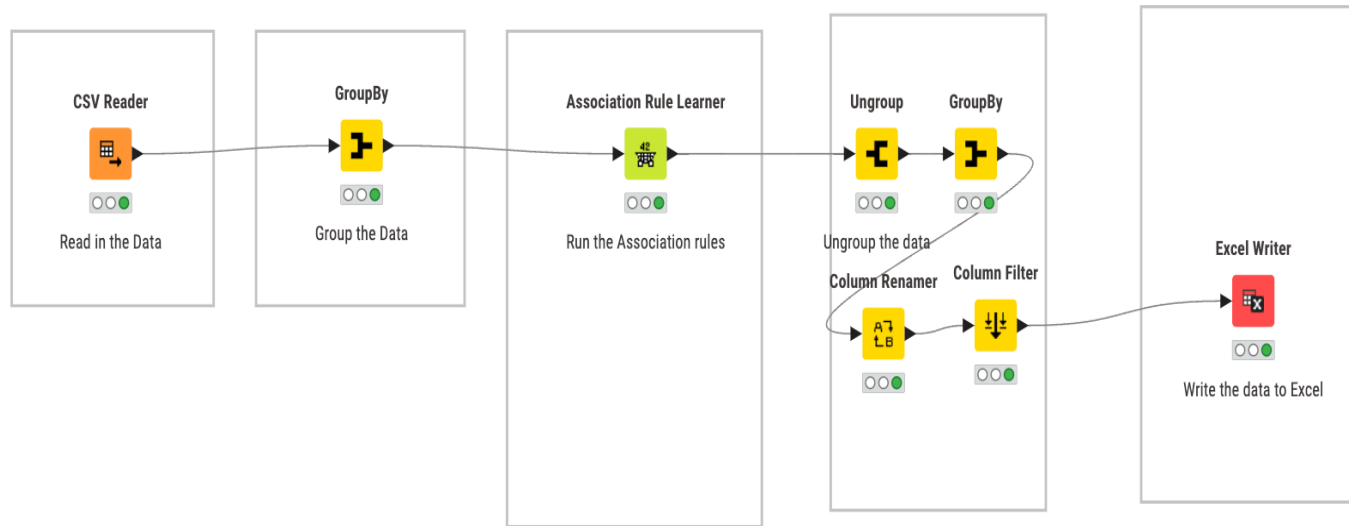
We have used 3 values to tune our KNIME Association Rule Learner.

Minimum support : 0.01 we have set this to get as man

Minimum confidence: 0.3

This is the minimum confidence that a rule should have

Maximum Itemset: this determines the maximum number of items in an item set. We set this to 10.



ASSOCIATION RULES

To identify useful rules in the data we look at rules that have high confidence and high lift. High confidence indicates a strong association between the itemset and the recommended item. High lift suggests that the presence of the itemset or basket significantly increases the likelihood of the recommended item being purchased which is what we want

Association Rules

Recommended	Support	Confidence	Lift	implies	Basket
flour	0.011414	1.000000	2.833333	<---	yogurt, all- purpose, pasta, individual meals, coffee/tea
fruits	0.011414	1.000000	2.699052	<---	butter, bagels, beef, mixes, soap
fruits	0.010536	1.000000	2.699052	<---	dishwashing liquid/detergent, beef, lunch meat, cereals, sugar, soap
fruits	0.010536	1.000000	2.699052	<---	dishwashing liquid/detergent, eggs, all- purpose, tortillas, soap
all- purpose	0.010536	1.000000	2.667447	<---	butter, fruits, beef, toilet paper, cereals, coffee/tea
all- purpose	0.011414	1.000000	2.667447	<---	dinner rolls, flour, milk, lunch meat, pork
all- purpose	0.011414	1.000000	2.667447	<---	shampoo, spaghetti sauce, toilet paper, coffee/tea, soap
all- purpose	0.010536	1.000000	2.667447	<---	yogurt, dishwashing liquid/detergent, flour, hand soap, laundry detergent, mixes
beef	0.010536	1.000000	2.667447	<---	eggs, spaghetti sauce, fruits, hand soap, sugar
beef	0.010536	1.000000	2.667447	<---	ice cream, hand soap, laundry detergent, tortillas, sugar
beef	0.010536	1.000000	2.667447	<---	poultry, fruits, milk, toilet paper, sugar, soap
beef	0.013169	1.000000	2.667447	<---	poultry, fruits, toilet paper, pork, sugar
beef	0.010536	1.000000	2.667447	<---	poultry, ice cream, cereals, juice, sandwich bags, aluminum foil
beef	0.010536	1.000000	2.667447	<---	shampoo, fruits, all- purpose, milk, coffee/tea
beef	0.010536	1.000000	2.667447	<---	shampoo, poultry, fruits, milk, cereals, aluminum foil
beef	0.010536	1.000000	2.667447	<---	shampoo, spaghetti sauce, poultry, fruits, toilet paper, aluminum foil

Top 5 unique product association rules based on lift and confidence.

1

FLOUR

Basket:Yoghurt, all-purpose,pasta,individual meals,coffee/tea

Recommendations:
flour

Support: 1.1%
Confidence: 1
Lift: 2.8

2

FRUITS

Basket:
Butter, bagels, beef, mixes, soap

Recommendations:
fruits

Support: 1.1%
Confidence: 1
Lift: 2.69

3

All-purpose

Basket:
butter, fruits, beef, toilet paper, cereals, coffee/tea

Recommendations:
All-purpose

Support: 1.1%
Confidence: 1
Lift: 2.67

4

All-purpose

Basket:
eggs, spaghetti sauce, fruits, hand soap, sugar

Recommendations:
Beef

Support: 1.1%
Confidence: 1
Lift: 2.67

Top 5 association rules based on support, lift and confidence.

5

C O F F E E / T E A

Basket:butter, ice cream,
waffles, tortillas, cereals

Recommendations:
Coffee/tea

Support: 1.1%
Confidence: 1
Lift: 2.64

Recommendations

We have identified a lot of different combinations in the data and associations. Based on the customer buying patterns and the market basket analysis we therefore make the following recommendations:

- **Bundle Offers:** For the strongest rules like eggs and paper towels, or dinner rolls and paper towels, consider creating bundle offers such as "Buy eggs and get 20% off on paper towels" or "Buy any two baked goods and get a paper towel roll at half price." These offers can encourage customers to increase their basket size.
- **Cross-Merchandising:** Position related items like beef and poultry or coffee and sandwich loaves in proximity within the store or in online categories to stimulate the purchasing of both.
- **Loyalty Rewards:** Offer extra loyalty points for buying combinations found in the top rules. For example, "Earn double points when you buy fruits and bagels together."
- **Targeted Discounts:** Provide targeted discounts via coupons or digital marketing for customers who have bought one item, to incentivize them to purchase the complementary item on their next visit.
- **Promotional Campaigns:** Use the insights from these rules to design promotional campaigns, highlighting the combination of items that have strong associations.
- **In-Store Signage:** Use in-store signage to highlight the product associations, especially for those with the highest lift and confidence such as a basket with beef and eggs, spaghetti sauce, fruits and other items to create a mental association.

Recommendations

We have identified a lot of different combinations in the data and associations. Based on the customer buying patterns and the market basket analysis we therefore make the following recommendations:

- **Buy two get one for free;** We can also stimulate extra sales by looking for combination items that are frequently bought together in pairs or more. This will help drive more sales and more output in terms of sales and these items would be paper towels, toilet paper, individual meals, milk, spaghetti sauce, fruits, waffles, ice cream, all-purpose
- **Marketing;** We suggest a robust marketing campaign to increase the number of orders per day in the store