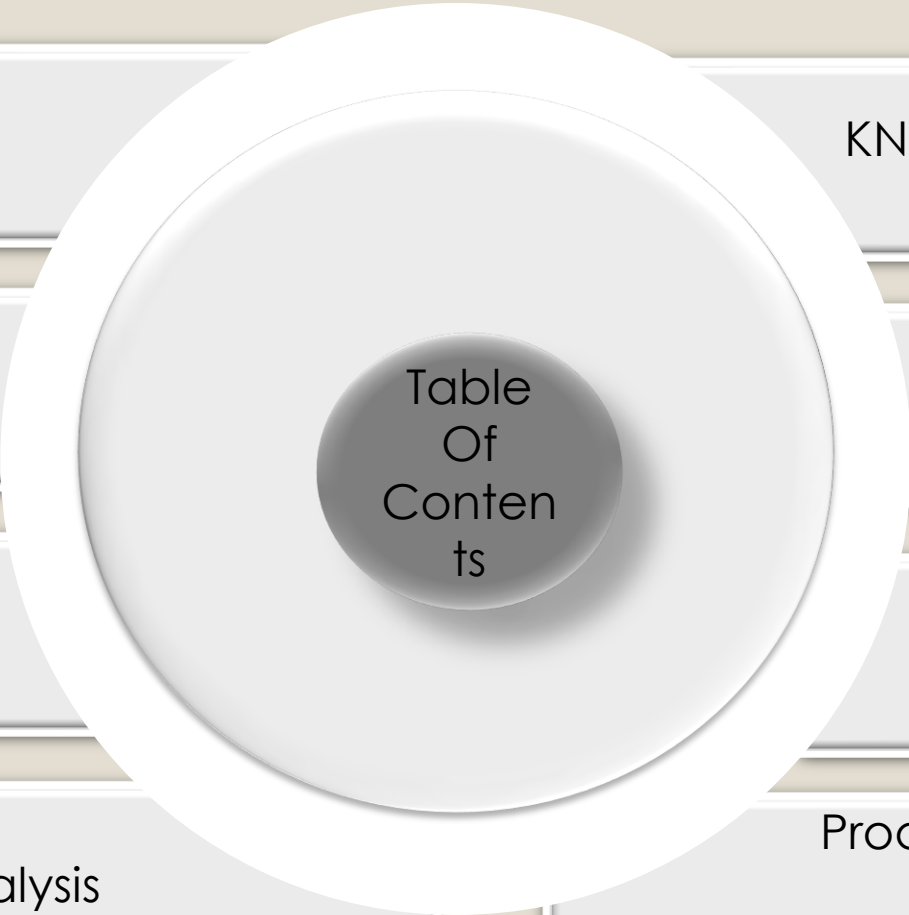# MARKET BASKET ANALYSIS

◦ MILESTONE 2 – DIVYA VIKRAM

Agenda

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.

Executive Summary of the data

We have 2 years and 2 months data of a Grocery store consisting of 20641 entries with 3 variables explaining the transaction ID, date of transaction and product information.

Table of Contents

Data Summary

- The dataset is Grocery store dataset with transactions spanning across 2 years and 2 months from Jan 2018 to Feb 2020.
- The data has 20641 entries and 3 columns. The data has 2 object data type and 1 integer data type.
- The dataset does not have null values.
- There are 4730 duplicates in the data. This is because same order ID has some items repeated.
- There are 1139 unique Order numbers and 37 unique products in the dataset .
- Poultry is the most frequently bought item.

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

Top 5 records in dataset

```
#   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)
```
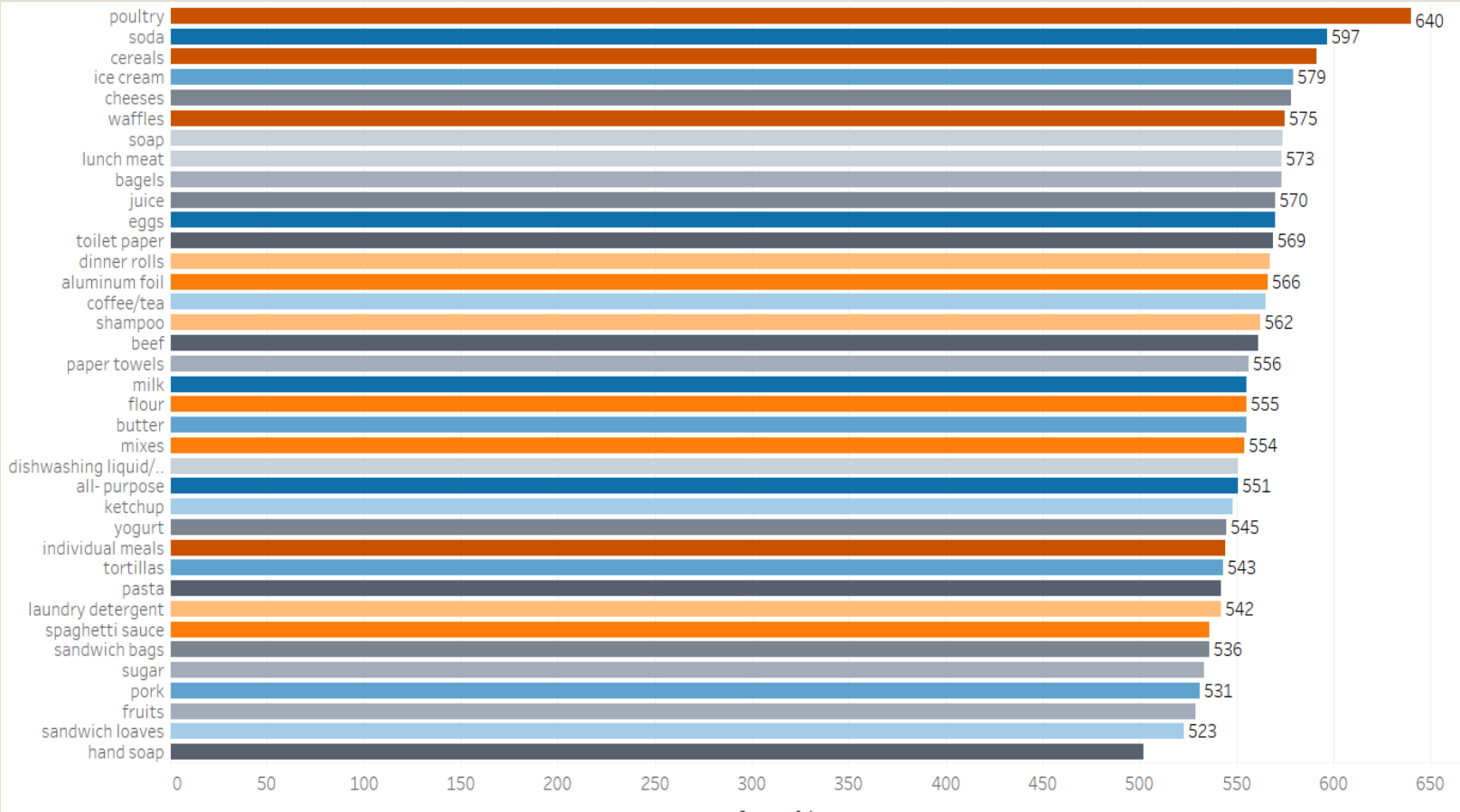
Columns Information

Data summary

	Date	Order_id	Product
count	20641	20641.000000	20641
unique	603	NaN	37
top	08-02-2019	NaN	poultry
freq	183	NaN	640
mean	NaN	575.986289	NaN
std	NaN	328.557078	NaN
min	NaN	1.000000	NaN
25%	NaN	292.000000	NaN
50%	NaN	581.000000	NaN
75%	NaN	862.000000	NaN
max	NaN	1139.000000	NaN

Exploratory Data Analysis

List of 37 Unique Items

all- purpose	
aluminum foil	
bagels	
beef	
butter	
cereals	
cheeses	
coffee/tea	
dinner rolls	
dishwashing liquid/detergent	
eggs	
flour	
fruits	
hand soap	
ice cream	
individual meals	
juice	
ketchup	
laundry detergent	
lunch meat	
milk	
mixes	
paper towels	
pasta	
pork	
poultry	
sandwich bags	
sandwich loaves	
shampoo	
soap	
soda	
spaghetti sauce	
sugar	
toilet paper	
tortillas	
waffles	
yogurt	

Items Count

| Item	Count
poultry	640
soda	597
cereals	
ice cream	579
cheeses	
waffles	575
soap	
lunch meat	573
bagels	
juice	570
eggs	
toilet paper	569
dinner rolls	
aluminum foil	566
coffee/tea	
shampoo	562
beef	
paper towels	556
milk	
flour	555
butter	
mixes	554
dishwashing liquid/..	
all- purpose	551
ketchup	
yogurt	545
individual meals	
tortillas	543
pasta	
laundry detergent	542
spaghetti sauce	
sandwich bags	536
sugar	
pork	531
fruits	
sandwich loaves	523
hand soap	

There are 37 Unique Items in the dataset
Poultry is the most popular Item with a count of 640
Hand soap is the least popular Item with a count of 502

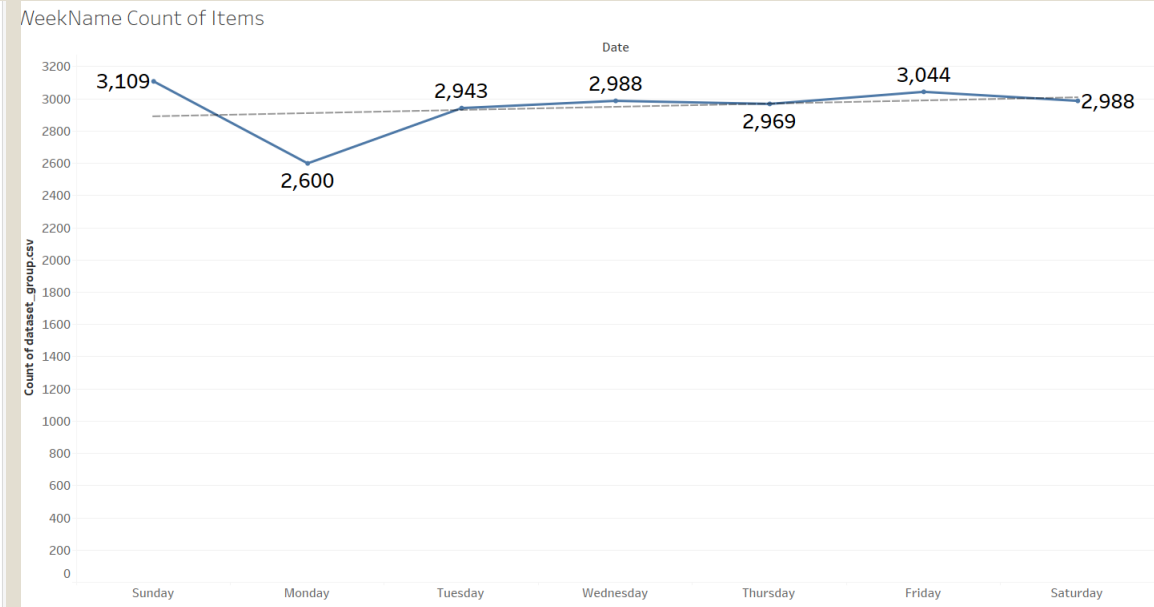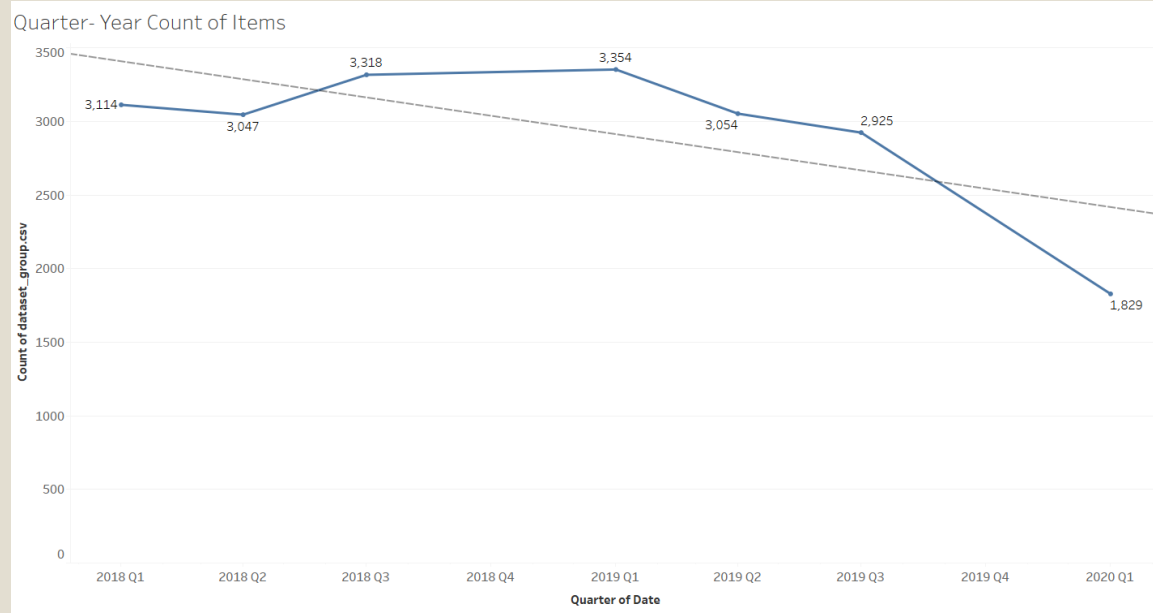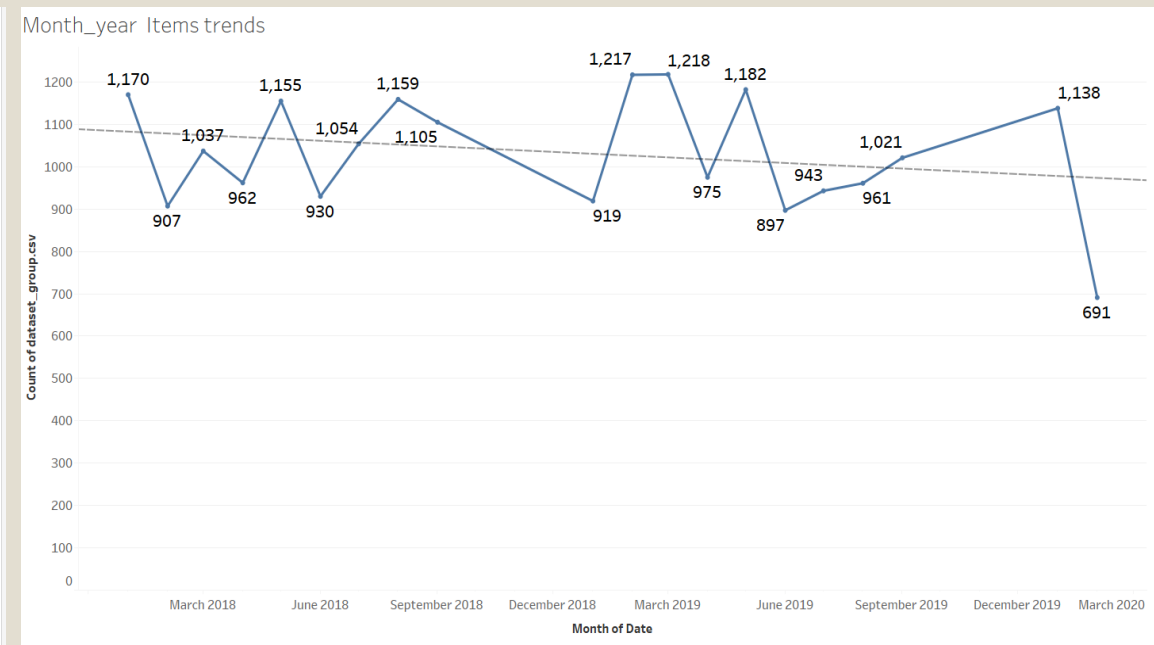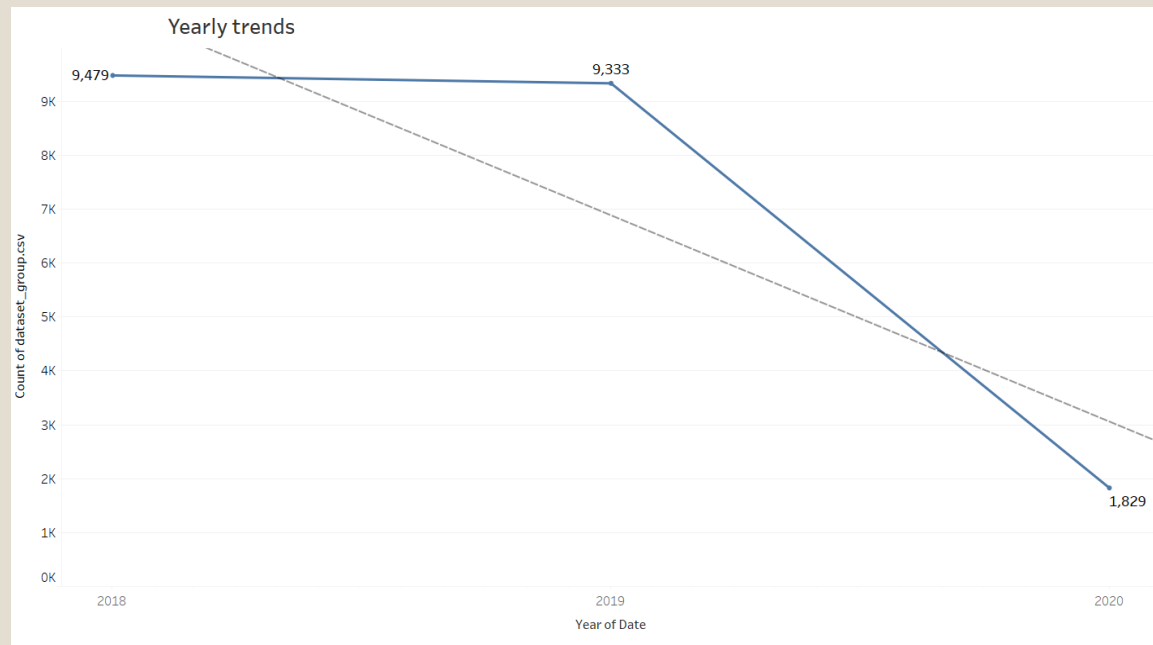**Number of Items per order**
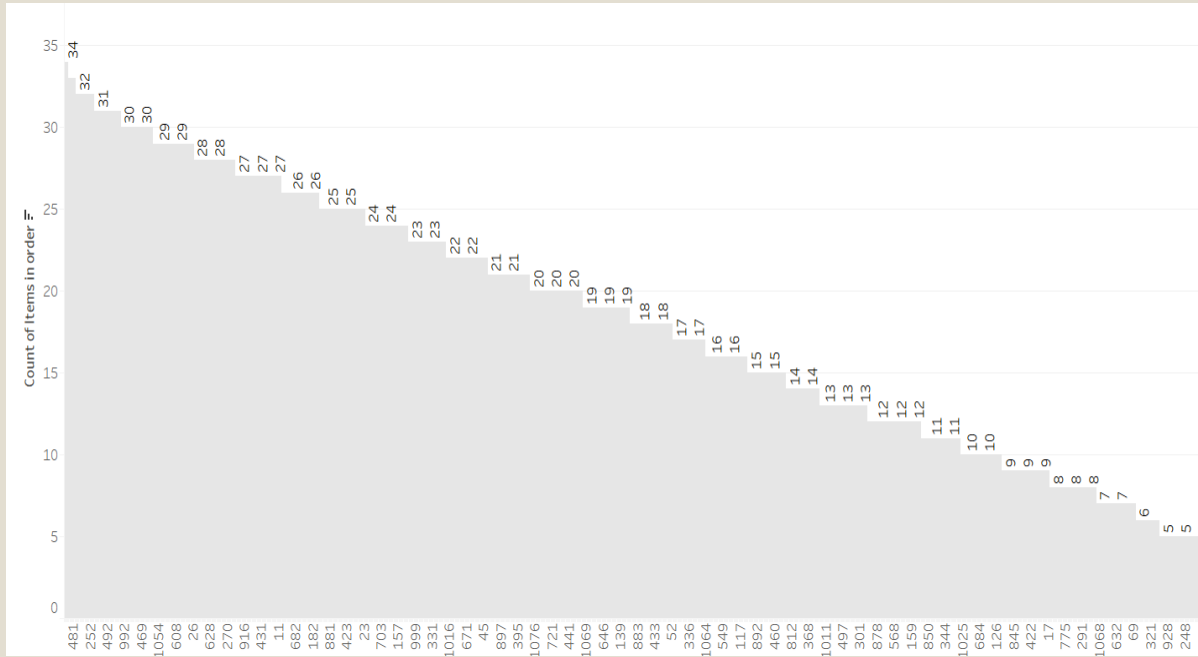
Number of Unique Items per order
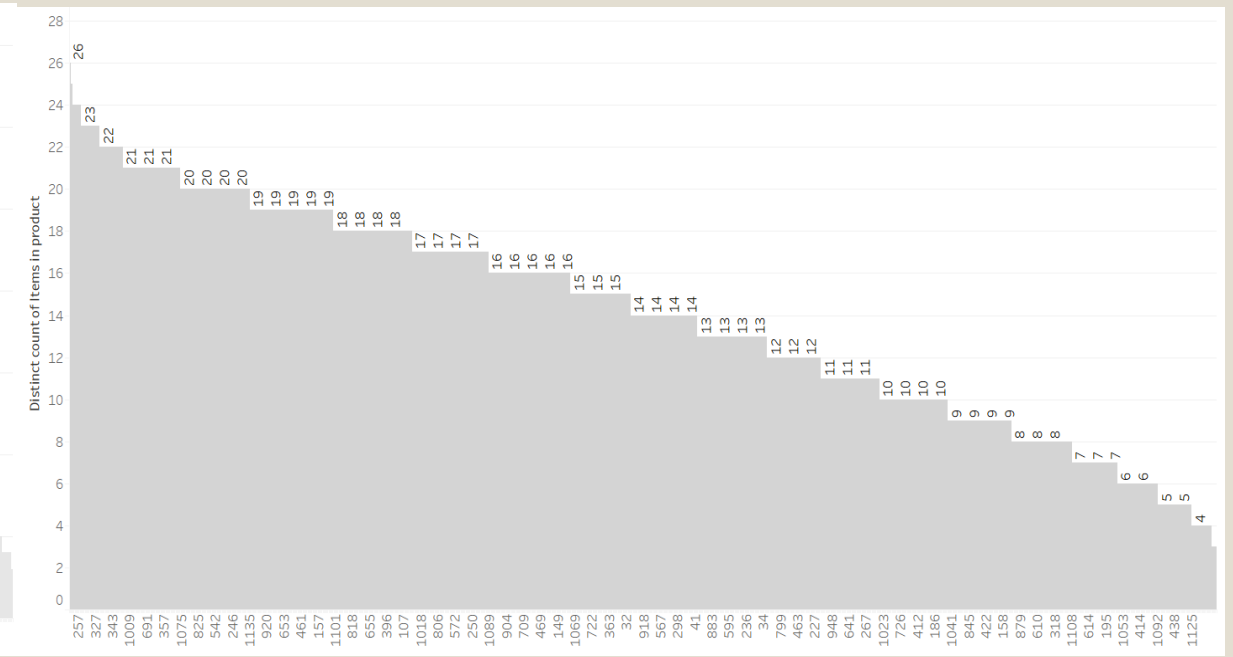
There are 1139 orders with each order having multiple items, sometimes items are repeating in order.

The above plots show the number of items In each order and also number of unique items in each order.

Maximum number on items in an order is 34 and minimum number of items in an order is 3
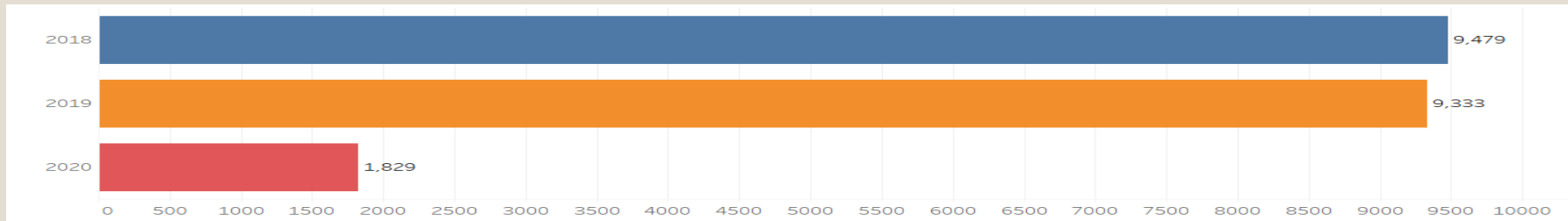
Maximum number on unique items in an order is 26 and minimum number of unique items in an order is 3

Sales Trends

Yearly trends

Count of dataset_group.csv — Year of Date

- 2018: 9,479
- 2019: 9,333
- 2020: 1,829

Month_year Items trends

Count of dataset_group.csv — Month of Date

1,170 · 907 · 1,037 · 962 · 1,155 · 930 · 1,054 · 1,159 · 1,105 · 919 · 1,217 · 1,218 · 975 · 1,182 · 897 · 943 · 961 · 1,021 · 1,138 · 691

Quarter- Year Count of Items

Count of dataset_group.csv — Quarter of Date

- 2018 Q1: 3,114
- 2018 Q2: 3,047
- 2018 Q3: 3,318
- 2019 Q1: 3,354
- 2019 Q2: 3,054
- 2019 Q3: 2,925
- 2020 Q1: 1,829

WeekName Count of Items

Date — Count of dataset_group.csv

- Sunday: 3,109
- Monday: 2,600
- Tuesday: 2,943
- Wednesday: 2,988
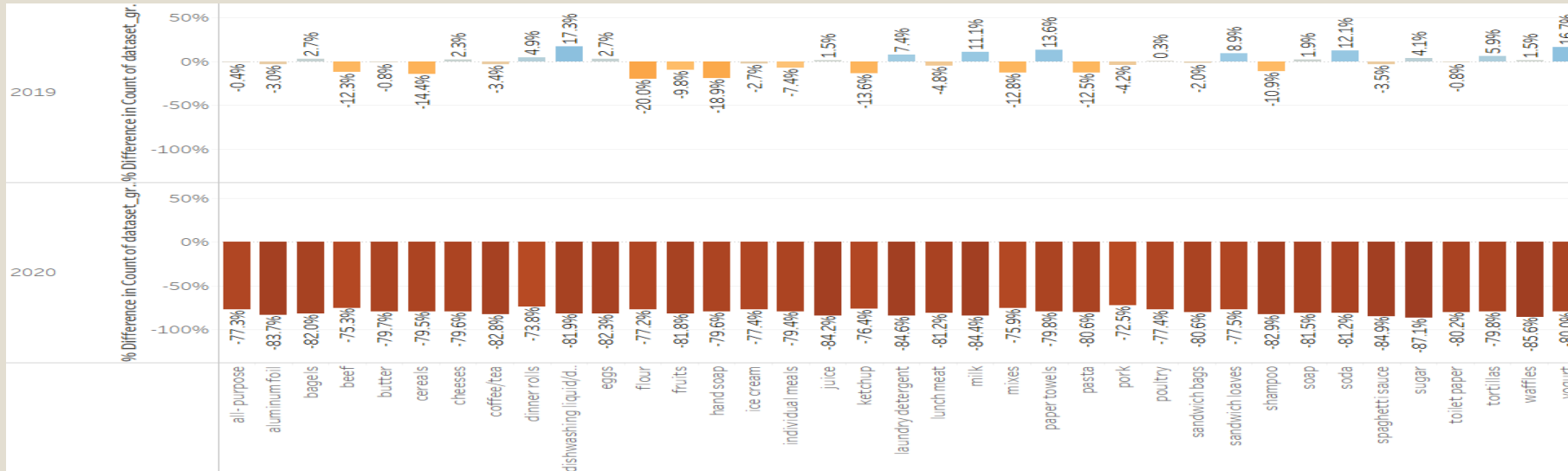- Thursday: 2,969
- Friday: 3,044
- Saturday: 2,988

- From yearly sales trend we can conclude that the number of items sold fell drastically from 2019 to 2020 .The highest number of items sold was in 2018 and then followed by 2019.
- From Quarterly sales trend we can see that The Q4 data has not been provided and 2020 has data only for Q1. On observing 3 quarters of 2018 and 2019 and 1 quarter of 2020, we can see that there is slight dip in sales in second quarters in both 2018 and 2019. However Q3 of 2018 , there was a increasing trend and Q3 of 2019, there was a decreasing trend.
- Monthly trends are not very conclusive and there is no seasonality , but there is sharp decline of 39% in items sold in January 2020.
- Daily Trends indicates Sunday has the highest items sold than any other weekday, Monday being the least , all other days there isn't much of a change

Number of Items sold Year on Year

2018: 9,479
2019: 9,333
2020: 1,829

Growth of number of Items sales Year on Year

2019 growth: all-purpose -0.4%, aluminum foil -3.0%, bagels 2.7%, beef -12.3%, butter -0.8%, cereals -14.4%, cheeses 2.3%, coffee/tea -3.4%, dinner rolls 4.9%, dishwashing liquid/d.. 17.3%, eggs 2.7%, flour -20.0%, fruits -9.8%, hand soap -18.9%, ice cream -2.7%, individual meals -7.4%, juice 1.5%, ketchup -13.6%, laundry detergent 7.4%, lunch meat -4.8%, milk 11.1%, mixes -12.8%, paper towels 13.6%, pasta -12.5%, pork -4.2%, poultry 0.3%, sandwich bags -2.0%, sandwich loaves 8.9%, shampoo -10.9%, soap 1.9%, soda 12.1%, spaghetti sauce -3.5%, sugar 4.1%, toilet paper -0.8%, tortillas 5.9%, waffles 1.5%, yogurt 16.7%

2020 growth: all-purpose -77.3%, aluminum foil -83.7%, bagels -82.0%, beef -75.3%, butter -79.7%, cereals -79.5%, cheeses -79.6%, coffee/tea -82.8%, dinner rolls -73.8%, dishwashing liquid/d.. -81.9%, eggs -82.3%, flour -77.2%, fruits -81.8%, hand soap -79.6%, ice cream -77.4%, individual meals -79.4%, juice -84.2%, ketchup -76.4%, laundry detergent -84.6%, lunch meat -81.2%, milk -84.4%, mixes -75.9%, paper towels -79.8%, pasta -80.6%, pork -72.5%, poultry -77.4%, sandwich bags -80.6%, sandwich loaves -77.5%, shampoo -82.9%, soap -81.5%, soda -81.2%, spaghetti sauce -84.9%, sugar -87.1%, toilet paper -80.2%, tortillas -79.8%, waffles -85.6%, yogurt -80.0%

Number of Items sold Year on Year

Product	2018	2019	2020
all- purpose	248	247	56
aluminum foil	266	258	42
bagels	259	266	48
beef	268	235	58
butter	253	251	51
cereals	291	249	51
cheeses	259	265	54
coffee/tea	265	256	44
dinner rolls	244	256	67
dishwashing liquid/..	231	271	49
eggs	258	265	47
flour	280	224	51
fruits	256	231	42
hand soap	254	206	42
ice cream	264	257	58
individual meals	257	238	49
juice	262	266	42
ketchup	265	229	54
laundry detergent	242	260	40
lunch meat	269	256	48
milk	243	270	42
mixes	266	232	56
paper towels	235	267	54
pasta	265	232	45
pork	239	229	63
poultry	287	288	65
sandwich bags	247	242	47
sandwich loaves	224	244	55
shampoo	275	245	42
soap	260	265	49
soda	256	287	54

- There were 9479 items in year 2018, 9333 in the year 2019 and 1829 items sold in 2020.
- The table on the right shows the number of times that product was ordered year on year.
- As we can see year 2020, all items have negative growth because we have data for only 2 months in 2020
- In year 2019 some items showed a positive growth , some showed a negative growth. Sales of dishwashing liquid has grown by 17.3% whereas Flour showed a negative growth of -20%

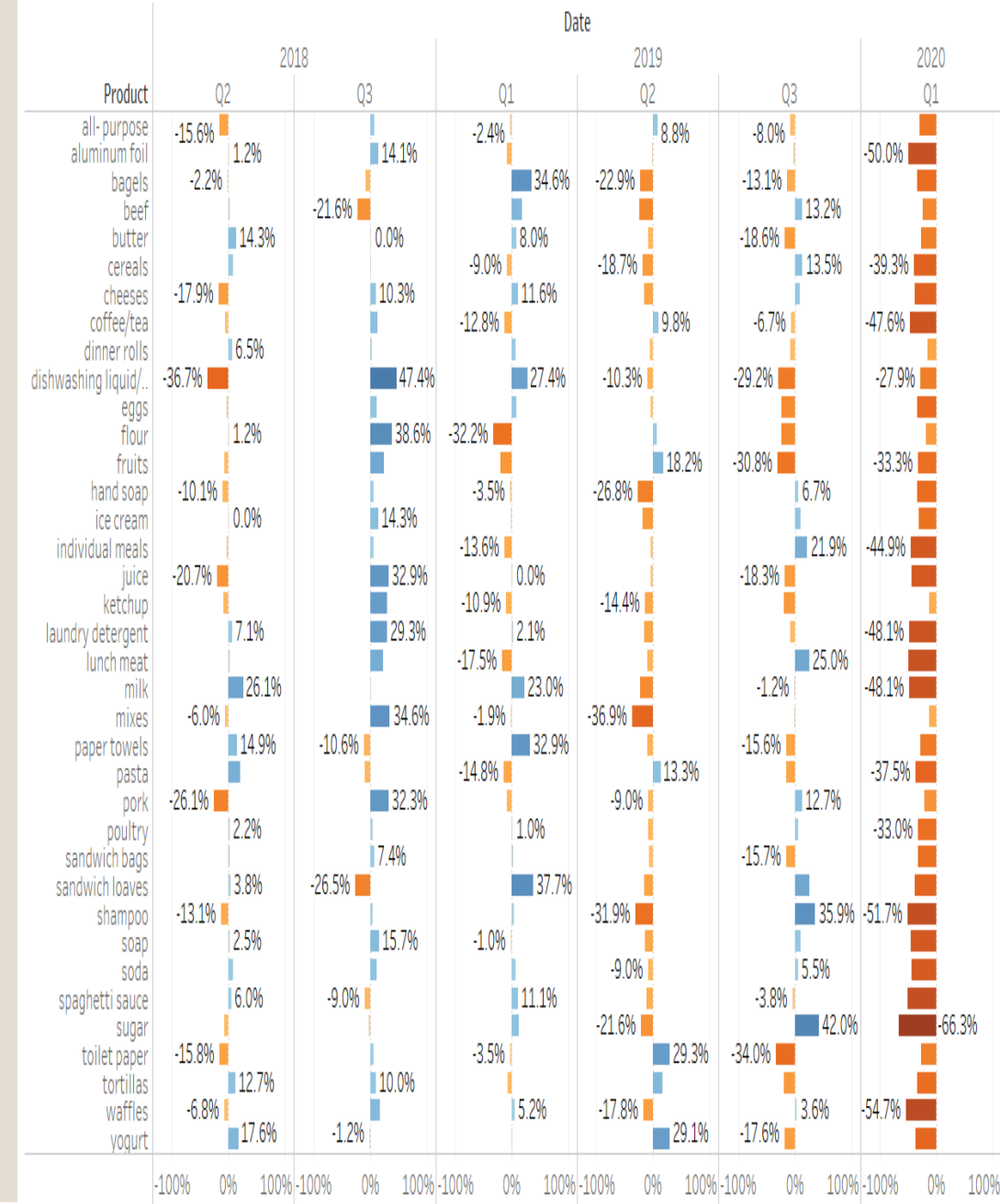## Number of Orders Year on Year

Distinct count of Order id

Year of Date		
2018		533
2019		507
2020		99
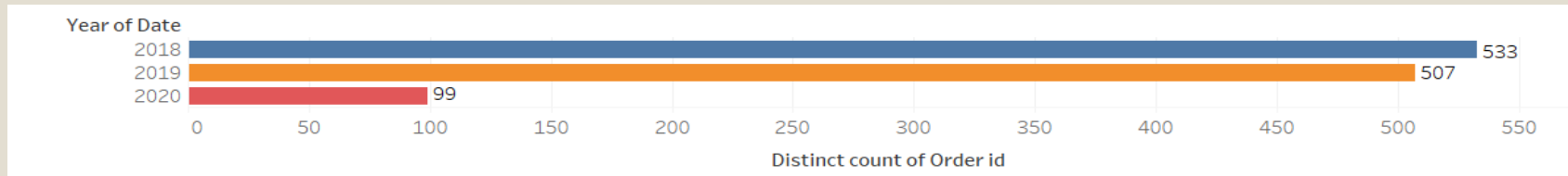
Year on Year growth in number of orders of the product

Product	Date 2018	2019	2020
all- purpose	198	190	39
aluminum foil	197	205	36
bagels	205	194	40
beef	203	182	42
butter	192	190	37
cereals	217	198	36
cheeses	203	201	41
coffee/tea	202	191	39
dinner rolls	195	203	45
dishwashing liquid/..	189	214	39
eggs	198	213	33
flour	202	165	35
fruits	207	185	30
hand soap	200	163	31
ice cream	207	206	41
individual meals	197	192	39
juice	196	201	32
ketchup	207	179	37
laundry detergent	195	201	35
lunch meat	210	203	37
milk	187	210	36
mixes	213	176	39
paper towels	179	198	36
pasta	207	180	36
pork	184	177	44
poultry	214	219	47
sandwich bags	191	187	41
sandwich loaves	178	184	36
shampoo	198	187	35
soap	192	199	41
soda	198	207	40

- There were 533 orders in year 2018, 507 in the year 2019 and 99 orders in 2020.
- The table on the right shows the number of orders of a certain product year on year.
- As we can see in year 2020, order on all items have negative growth because we have data for only 2 months in 2020
- In year 2019 order of some items showed a positive growth , some showed a negative growth. Orders of dishwashing liquid has grown by 13.23% whereas Handsoap showed a negative growth of -18.5 % per order

Number of Items sold Quarter on Quarter

Growth of number of Items sales Quarter on Quarter

Year of D..	Quarter ..		
2018	Q1		3,114
	Q2		3,047
	Q3		3,318
2019	Q1		3,354
	Q2		3,054
	Q3		2,925
2020	Q1		1,829

- Maximum number of Items ordered were in Quarter 1 of 2019 in which 3354 Items were ordered
- Least number of Items ordered were in Quarter 1 of 2020 in which 1829 Items were ordered
- As we can see Q 1 of year 2020, all items have negative growth because we have data for only 2 months in 2020
- Quarter on Quarter growth of each Items in Year 2018 and 2019 indicates Q3 of year 2019 had highest growth of 47.4% for Dishwashing Liquid but interestingly the least growth of -36.7% was in Q2 of 2018 for Dishwashing Liquid

Date

Product	2018 Q2	2018 Q3	2019 Q1	2019 Q2	2019 Q3	2020 Q1
all- purpose	-15.6%		-2.4%	8.8%	-8.0%	
aluminum foil	1.2%	14.1%				-50.0%
bagels	-2.2%		34.6%	-22.9%	-13.1%	
beef		-21.6%				13.2%
butter	14.3%	0.0%	8.0%		-18.6%	
cereals			-9.0%	-18.7%	13.5%	-39.3%
cheeses	-17.9%	10.3%	11.6%			
coffee/tea			-12.8%	9.8%	-6.7%	-47.6%
dinner rolls	6.5%					
dishwashing liquid/..	-36.7%	47.4%	27.4%	-10.3%	-29.2%	-27.9%
eggs						
flour	1.2%	38.6%	-32.2%			
fruits				18.2%	-30.8%	-33.3%
hand soap	-10.1%		-3.5%	-26.8%	6.7%	
ice cream	0.0%	14.3%				
individual meals			-13.6%		21.9%	-44.9%
juice	-20.7%	32.9%	0.0%		-18.3%	
ketchup			-10.9%	-14.4%		-48.1%
laundry detergent	7.1%	29.3%	2.1%		25.0%	
lunch meat			-17.5%			-48.1%
milk	26.1%		23.0%		-1.2%	
mixes	-6.0%	34.6%	-1.9%	-36.9%	-15.6%	
paper towels	14.9%	-10.6%		32.9%		-37.5%
pasta			-14.8%	13.3%		
pork	-26.1%	32.3%		-9.0%	12.7%	-33.0%
poultry	2.2%		1.0%		-15.7%	
sandwich bags		7.4%				
sandwich loaves	3.8%	-26.5%	37.7%		35.9%	-51.7%
shampoo	-13.1%			-31.9%		
soap	2.5%	15.7%	-1.0%		5.5%	
soda				-9.0%	-3.8%	
spaghetti sauce	6.0%	-9.0%	11.1%		42.0%	-66.3%
sugar				-21.6%		
toilet paper	-15.8%		-3.5%	29.3%	-34.0%	
tortillas	12.7%	10.0%			3.6%	-54.7%
waffles	-6.8%		5.2%	-17.8%		
yogurt	17.6%	-1.2%		29.1%	-17.6%	

Number of Orders per Quarter
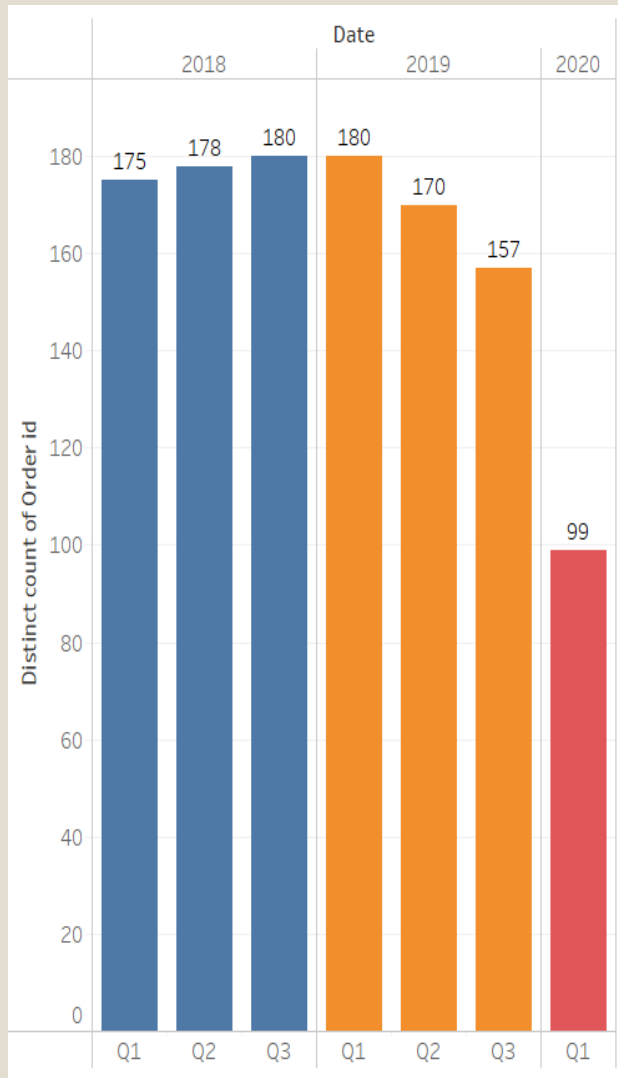
Quarter on quarter growth in number of orders of the product
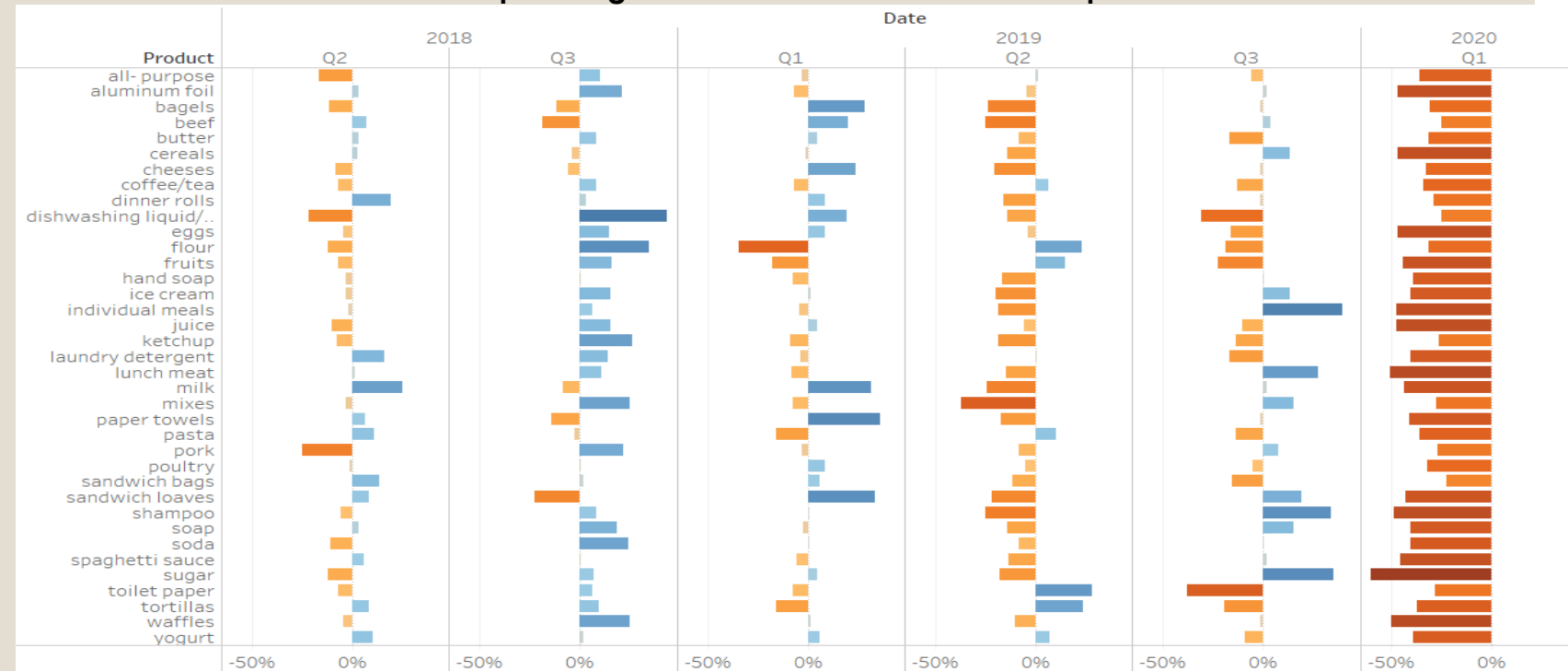
- We have data for only 3 quarters for year 2018 and 2019 and one quarter for 2020. Q3 of 2018 and Q1 of 2019 had highest number of orders of 180 each Q1 of 2020 had least number of orders with just 99.
- Q3 of 2018 and Q1 of 2019 had fewer negative growths as we can see from the plot above and Q1 of 2020 had negative growth of all products orders
- Quarter on Quarter growth of products in orders in Year 2018 and 2019 indicates that highest growth was that of dishwashing liquid with an increase of 43.1% in Q3 of 2018 and lowest was that of toilet paper which had a negative growth of -38.3% in quarter Q3 of 2019

Number of Orders per Month
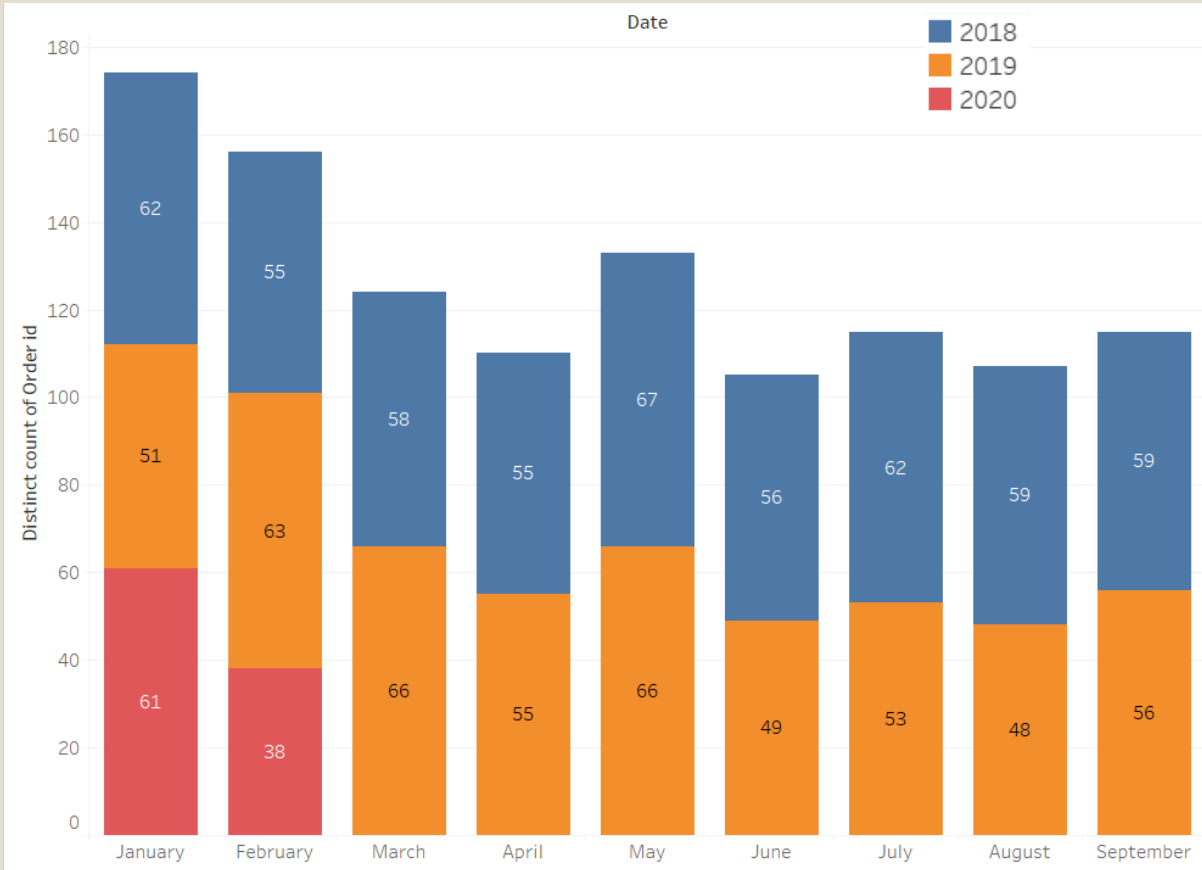
Number of Items sold per Month
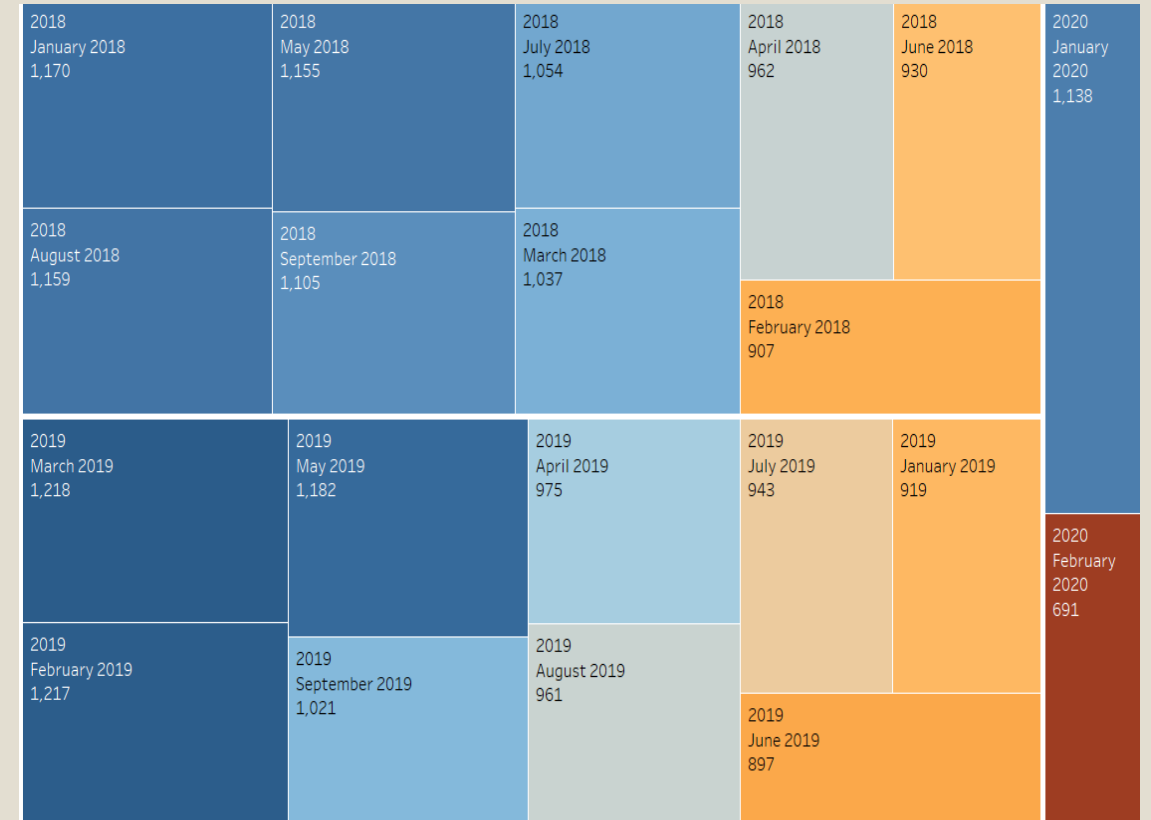
January had the highest orders and highest number of items sold across all years and June had least number of orders

Monthly Analysis indicates that May 2018 had the highest number of orders of 67 and February 2020 had the least orders of 38.

March 2019 had the highest number of items sold of 1218 items and February 2020 had the least number of items ordered of 691.

Inference from EDA

- Poultry is the most product followed by soda, ice-cream and cereals and cheeses
- Hand Soap is a product least ordered product, followed by Sandwich loaves and fruits, pork and sugar.
- The orders decreasing over the years with 2018 receiving highest followed closely by 2019 and 2020 had least number of orders .
- In January months highest numbers of orders are placed.
- Sunday seems to be receiving highest number of transactions
- Top 3 products ordered in 2018 are Cereals, Poultry & Flour
- Top 3 products ordered in 2019 are Poultry, Soda & Dishwashing liquid/Detergent
- Top 3 products ordered in 2020 are Dinner rolls , Poultry and pork

Market Basket Analysis

◦ **Market basket analysis(MBA)** is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns . It also helps in increasing customer engagement ,optimizing marketing strategies. It involves analysing large data sets, such as purchase history, to reveal product groupings, as well as products that are likely to be purchased together.

 Association Rules in market basket analysis

◦ In market basket analysis, association rules are used to predict the likelihood of products being purchased together. Association rules count the frequency of items that occur together, seeking to find associations that occur far more often than expected. Using customer purchase history data retailers can determine that products are often purchased together, retailers can offer special deals , they can optimize product placement, or create new product bundles to encourage further sales of these combinations. This way market basket analysis can help increase the sales and customer satisfaction by making shopping experience more productive and valuable for customers. With MBA retailer can make targeted marketing campaign/sales promotion . It also help is designing store layout based on customer behavior patterns . By using market basket analysis, customers may feel a stronger sentiment or brand loyalty towards the company.

◦ The most well-known case study using market basket analysis is probably Amazon.com. As soon as you visit Amazon to look at a product, the product description will suggest "Items purchased together frequently." It is the an example of Market Basket Analysis cross-selling tactics.

◦ To carry out an MBA you'll first need a data set of transactions , in our case we have 1139 transactions .

◦ Each row/transaction represents a group of items or products that have been bought together and often referred to as an "itemset".

◦ For example, one itemset might be: {cheese, cereals, sugar, bagels and soda} in which case all of these items have been bought in a single transaction.

◦ In an MBA, the transactions are analyzed to identify rules of association. For example, one rule could be:

{fruits, juice} => {Bagels}.

The set of items on the left-hand side (**fruits, juice)** in the example above) is the antecedent of the rule, while the one to the right (**Bagels**) is the consequent

◦ This means that if a customer has a transaction that contains fruits and juice, then they are likely to be interested in also buying Bagels.

◦ Before acting on a rule, a retailer needs to know whether there is sufficient evidence to suggest that it will result in a beneficial outcome. We therefore measure the strength of a rule by calculating the following three commonly used metrics:

❖ Support

❖ Confidence

❖ Lift

Dataset with Order ID and Products in the basket

◦ **Support :** Its the percentage of transactions that contain all of the items in an itemset (e.g., Fruits, Juice and Bagel). The higher the support the more frequently the itemset occurs. Rules with a high support are preferred since they are likely to be applicable to a large number of future transactions.

If there is rule $X => Y$, meaning if X then Y

Support(X) = (Number of transactions containing X) / (Total number of transactions)

where X is the itemset for which you are calculating the support.

◦ **Confidence:** Its the probability that a transaction that contains the items on the left hand side of the rule (in our example, Fruits and Juice) also contains the item on the right hand side (Bagels). The higher the confidence, the greater the likelihood that the item on the right hand side will be purchased or, in other words, the greater the return rate you can expect for a given rule.

Confidence(X => Y) = (Number of transactions containing X and Y) / (Number of transactions containing X)

Or

Confidence(X => Y)   = Supp(X U Y)/Supp(X)

Lift: Its the probability of all of the items in a rule occurring together (otherwise known as the support) divided by   the product of the probabilities of the items on the left and right hand side occurring as if there was no association between them.

Lift(X=>Y) = Supp(X U Y)/Supp(Y)*Supp(X)

or

Lift(X=>Y) = Conf(X=>Y)/Supp(Y)

◦ Lift value near 1 indicates X and Y almost often appear together as expected, greater than 1 means they appear together more than expected and less than 1 means they appear less than expected. Greater lift values indicate stronger association.

◦ As an example, if Fruits, Juice and Bagel occurred together in 2.5% of all transactions, Fruits and Juice in 10% of transactions and Bagel in 8% of transactions, then the lift would be: 0.025/(0.1*0.08) = 3.125. A lift of more than 1 suggests that the presence of Fruits and Juice increases the probability that a Bagel will also occur in the transaction. Overall, lift summarizes the strength of association between the products on the left and right hand side of the rule; the larger the lift the greater the link between the two products.

• Support and Confidence is set by the minimum support and minimum confidence thresholds. These thresholds help to compare the rule strength according to your own will. The closer to threshold the more the rule is of use to the client.

• When you apply Association Rule Mining on a given set of transactions T 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

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

KNIME workflow

We are using KNIME tool to create different association rules for the dataset .
- First Step is group the Items based on Order ID using GroupBy Node
- Second Step is to create Item sets for an order by removing duplicate items in an order using Cell Splitter node.
- Third step is create association rules by providing the threshold values for Support and confidence using Association rule learner node.
- We run the Association Rule Learner node on a set of basket, to generate the association rule set. Each rule includes a collection of product IDs as antecedent, one product ID as consequent, and quality measures, such as support, confidence, and lift. Quality measures describe how likely and how reliable the current rule is. You can make your association rule engine larger or smaller, restrictive or tolerant, by changing the threshold values in the Association Rule Learner configuration settings, such as "minimum set size", "minimum rule confidence", and "minimum support".
- The figure below is KNIME workflow Image
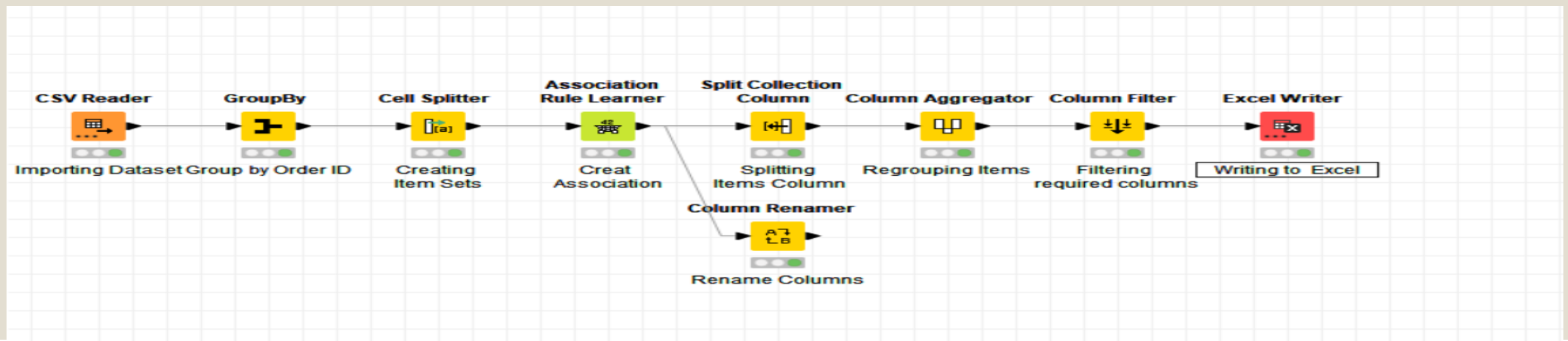
Support ,Confidence and Lift Thresholds

- There is no particular threshold value of support. It can be set based on how many rules of association we need, more the threshold, lesser the number of rules generated.
- Since we have many items in an order its advisable to generate rule contains multiple items in the antecedent. For our analysis
- The Minimum support is set at 0.05.
- The Minimum confidence is at 0.1
- Maximal Itemset Length is 10

- The associations with highest lift values are considered to be accepted more.
- With these threshold values 24600 rules are generated.

Output Image and Associations

Association rule in Tabular form

able "default" - Rows: 24600   Spec - Columns: 6   Properties   Flow Variables

Row ID	D Support	D Confide...	D ▼ Lift	S Recommended Item	S implies	[...] Basket Items
rule1142	0.055	0.649	1.791	paper towels	<---	[eggs,ice cream,pasta]
rule1141	0.055	0.643	1.731	pasta	<---	[paper towels,eggs,ice cream]
rule138	0.051	0.674	1.726	cheeses	<---	[bagels,cereals,sandwich bags]
rule51	0.05	0.64	1.7	juice	<---	[yogurt,toilet paper,aluminum foil]
rule135	0.051	0.63	1.678	mixes	<---	[yogurt,poultry,aluminum foil]
rule137	0.051	0.611	1.66	sandwich bags	<---	[cheeses,bagels,cereals]
rule616	0.054	0.642	1.651	dinner rolls	<---	[spaghetti sauce,poultry,laundry detergent]
rule289	0.052	0.641	1.649	dinner rolls	<---	[spaghetti sauce,poultry,ice cream]
rule55	0.05	0.62	1.645	juice	<---	[yogurt,poultry,aluminum foil]
rule292	0.052	0.686	1.628	poultry	<---	[dinner rolls,spaghetti sauce,ice cream]
rule298	0.052	0.634	1.627	eggs	<---	[paper towels,dinner rolls,pasta]
rule299	0.052	0.602	1.621	pasta	<---	[paper towels,eggs,dinner rolls]
rule141	0.051	0.63	1.621	dinner rolls	<---	[spaghetti sauce,poultry,cereals]
rule1140	0.055	0.63	1.616	eggs	<---	[paper towels,ice cream,pasta]
rule59	0.05	0.613	1.616	coffee/tea	<---	[yogurt,cheeses,cereals]
rule293	0.052	0.628	1.614	dinner rolls	<---	[spaghetti sauce,poultry,juice]
rule284	0.052	0.628	1.61	eggs	<---	[dinner rolls,poultry,soda]
rule618	0.054	0.598	1.603	spaghetti sauce	<---	[dinner rolls,poultry,laundry detergent]
rule146	0.051	0.604	1.589	milk	<---	[poultry,laundry detergent,cereals]
rule291	0.052	0.59	1.581	spaghetti sauce	<---	[dinner rolls,poultry,ice cream]
rule294	0.052	0.584	1.566	spaghetti sauce	<---	[dinner rolls,poultry,juice]
rule1139	0.055	0.624	1.565	ice cream	<---	[paper towels,eggs,pasta]
rule300	0.052	0.567	1.565	paper towels	<---	[eggs,dinner rolls,pasta]
rule60	0.05	0.588	1.564	mixes	<---	[dishwashing liquid/detergent,poultry,laundry detergent]
rule139	0.051	0.617	1.558	cereals	<---	[cheeses,bagels,sandwich bags]
rule619	0.054	0.656	1.556	poultry	<---	[dinner rolls,spaghetti sauce,laundry detergent]
rule49	0.05	0.594	1.544	aluminum foil	<---	[yogurt,toilet paper,juice]
rule56	0.05	0.588	1.528	yogurt	<---	[cheeses,cereals,coffee/tea]
rule57	0.05	0.594	1.52	cheeses	<---	[yogurt,cereals,coffee/tea]
rule145	0.051	0.574	1.518	laundry detergent	<---	[poultry,milk,cereals]
rule144	0.051	0.637	1.512	poultry	<---	[dinner rolls,spaghetti sauce,cereals]
rule140	0.051	0.58	1.505	bagels	<---	[cheeses,cereals,sandwich bags]
rule297	0.052	0.584	1.502	dinner rolls	<---	[paper towels,eggs,pasta]
rule22522	0.083	0.563	1.498	individual meals	<---	[sandwich loaves,lunch meat]
rule48	0.05	0.576	1.497	yogurt	<---	[toilet paper,juice,aluminum foil]
rule23265	0.099	0.579	1.49	dinner rolls	<---	[spaghetti sauce,poultry]
rule286	0.052	0.578	1.487	dinner rolls	<---	[eggs,pasta,soda]
rule21203	0.078	0.56	1.486	juice	<---	[shampoo,spaghetti sauce]
rule62	0.05	0.576	1.484	dishwashing liquid/detergent	<---	[poultry,laundry detergent,mixes]
rule143	0.051	0.552	1.48	spaghetti sauce	<---	[dinner rolls,poultry,cereals]
rule134	0.051	0.569	1.479	aluminum foil	<---	[yogurt,poultry,mixes]
rule288	0.052	0.546	1.471	pasta	<---	[eggs,dinner rolls,soda]
rule22628	0.083	0.514	1.47	sandwich loaves	<---	[cheeses,ketchup]
rule287	0.052	0.573	1.469	eggs	<---	[dinner rolls,pasta,soda]
rule58	0.05	0.582	1.469	cereals	<---	[yogurt,cheeses,coffee/tea]

- The association rule is formed with highest lift value that is 1.791 is at a confidence level of 0.649 and support of 0.055.This rule implies that the customers who bought Eggs, Ice cream and Pasta are most likely to buy paper towels. The larger lift indicates the greater the link between the two products
- The association rule formed with highest support value that is 0.195 is at a confidence level of 0.462 and lift of 1.189.This rule implies that the customers who bought Poultry are most likely to buy Dinner Rolls . Rules with a high support are preferred since they are likely to be applicable to a large number of future transactions.
- The association rule is formed with highest confidence value that is 0.686 is at a support level of 0.052 and lift of 1.628.This rule implies that the customers who bought Dinner Rolls, spaghetti sauce , ice cream are most likely to buy Poultry .Higher the confidence the greater the likelihood that the recommended item will be purchased or, in other words, the greater the return rate you can expect for a given rule.

Insights and Business Recommendations

The sales have been dropping year on year . The widespread changes in grocery shopping behavior could be due to online shopping, pandemic etc as people have gotten addicted to the time-savings and ease of using door to door delivery to do full grocery runs — or to get a handful of items needed to make a special meal without having to leave the house and also shop online to save money and find great deals. Other factors for dropping orders could be poor customer service , economic instability, inflation, and an out-of-control cost of living crisis. In order to increase sales, stores must be ready to be flexible, to provide attractive deals and creative offerings, create grocery ads informed by their community's needs, and meet their customers where they are. Here are some ways to increase grocery sales, even during challenging times.

◦ Creating a marketing plan and build the brand on social media. From email marketing to social media to video marketing, there are many tools available to grocery businesses.

◦ Create and share effective supermarket promotions based on demographics, desires, socio-economic data, such as age, gender, ethnicity, and income and needs of your community, and build your grocery promotions from there. Classic promotions like 0 delivery charge on online delivery , weekly flyer specials, and limited-time products for holidays could be options . This helps to develop marketing strategies that break customers into segments and designs marketing activities that target the segments most likely to respond to your efforts

◦ See if there are any sticking points or roadblocks, if there's a pileup of people during peak hours, or if some products should be moved to a better location for optimal sales. For example, store may typically keep butter in a totally different area of the store as fresh corn, but you can test out putting a box of corn near the dairy aisle at the height of corn grilling season — with a vibrant display that invites shoppers to consider grilling tonight.

◦ Get creative with grocery item bundles. A great way to boost sales is pairing products together for an appealing price, like offering butter, eggs, all-purpose flours, sugar or creating baking kits for popular holiday desserts like pumpkin pie. With bundles you can play with pricing to still appeal to the customer, but increase your overall profit margin. Tailoring your bundles to the needs of your community ensures they'll be successful. For example, if you have a lot of young people in your area who prefer quick meals , consider quick meal bundles like individual meals, waffles , fruits etc .

◦ Understanding items that are frequently bought together can help you manage your inventory better. For example, you can avoid stockouts when you run a promotion by predicting how the sale of products related to the promotional item may increase.

◦ Selling low-margin products can drive the sale of high-margin products. Understanding their associations can help you develop loss-leader strategies with a low-priced product that increases store and website traffic, and increases overall revenue and profits.

◦ Identifying products that are often purchased together through market basket analysis helps optimize the placement of products in stores, in catalogs and on a web site. Placing those associated items close by increases encourages the buyer to purchase multiple products

Product Recommendations and Offers

- As per data analysis Poultry , Soda, Cereals 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 of variety to choose from .

- Hand soaps and Sandwich loaves are least purchased items. So offering them on discount or as a free item as part of bundle or on shopping above a threshold amount might increase sales.

- Sale of dishwashing liquid increased by 17% in 2019, so it seems like popular item and as data is available only for 2 months in 2020, there is a chance that it might grow further ,

- Demand of hand soap fell by 18.5% in 2019, may be offering it as combo of hygiene pack might increase sales or giving offers like buy 1 get 1 free might increase the sales further

- People who buy eggs, ice cream, pasta will most likely buy paper towels as these associations have a higher lift , they can be bundled as combo pack . Bagel, cereals , sandwich bags and cheeses have a very high lift, therefore store can offer a combo pack of Bagel, cereals ,sandwich bags and cheeses.

- Also since Eggs, ice-cream , pasta are most preferred basket items, different flavours of ice-cream , different kinds of eggs and varieties of pasta can be on offer for customers to choose from

- Poultry , Yogurt, dinner rolls are the most recommended items, they can be offered at a discount or as free item on total bill amount above a certain threshold . Stocking sufficient quantities so that they are available at all the times also helps.

- If not purchased as a separate item, Poultry , Yogurt, dinner rolls can be made available to the customers at discounted prices.

- Poultry with Mixes also has a high lift and can be sold as a lunch or dinner combo as well.
- Poultry with dinner rolls, spaghetti sauce and ice cream also has a high lift and can be sold as a lunch or dinner combo as well.
- Sandwich loaves is also one of the least selling products. From the output table we see that people who buy loaves are most likely buy individual meals. So the store can provide combo offer of buy 2 meals and get 1 sandwich loaf free. It may be placed in the racks around the sandwich loaves and may be sold at a discount
- Different flavours of Yogurt and different verities of Poultry and   dinner rolls can be offered since it's the most recommended items
- Sales of hand soaps is not very high , therefore they can be bundled as a combo item along with dishwashing liquid/detergent to increase the sales as the lift of the combo is quite high.
- A special discount coupon on the next shopping on purchasing a combo like eggs, ice cream, pasta or Bagel, cereals and cheeses can be offered to customers so that they will buy the combo even though they may not intend to buy one or more items in that combo at that moment .
- A 5% off on current bill or the least priced item waived off which ever is lesser kind of offer on lesser sold items .If shopped over a certain threshold amount offer to give 5 % discount, and place the least sold item near counter , so that customers can just add up to the bill by just buying the items near the counter , just to get that discount

- Poultry with Mixes also has a high lift and can be sold as a lunch or dinner combo as well.
- Poultry with dinner rolls, spaghetti sauce and ice cream also has a high lift and can be sold as a lunch or dinner combo as well.
- Sandwich loaves is also one of the least selling products. From the output table we see that people who buy loaves are most likely buy individual meals. So the store can provide combo offer of buy 2 meals and get 1 sandwich loaf free. It may be placed in the racks around the sandwich loaves and may be sold at a discount
- Different flavours of Yogurt and different verities of Poultry and   dinner rolls can be offered since it's the most recommended items
- Sales of hand soaps is not very high , therefore they can be bundled as a combo item along with dishwashing liquid/detergent to increase the sales as the lift of the combo is quite high.
- A special discount coupon on the next shopping on purchasing a combo like eggs, ice cream, pasta or Bagel, cereals and cheeses can be offered to customers so that they will buy the combo even though they may not intend to buy one or more items in that combo at that moment .
- A 5% off on current bill or the least priced item waived off which ever is lesser kind of offer on lesser sold items .If shopped over a certain threshold amount offer to give 5 % discount, and place the least sold item near counter , so that customers can just add up to the bill by just buying the items near the counter , just to get that discount