

# Marketing Retail Analytics

Bhushan Rai  
Great Learning  
PGDSBA2002-203



# Contents:

1. **:Exploratory Analysis --> Exploratory Analysis of data & an executive summary,,**
2. **Trends across months/years/quarters/days**
3. **Use of Market Basket Analysis (Association Rules)**
4. **KNIME workflow image**
5. **Threshold values of Support and Confidence**
6. **Associations Identified -confidence, & lift values**
7. **Suggested Possible Combos and Lucrative Offers**

# Business 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
- ❖ Data Dictionary:  
Date, Order ID, Product

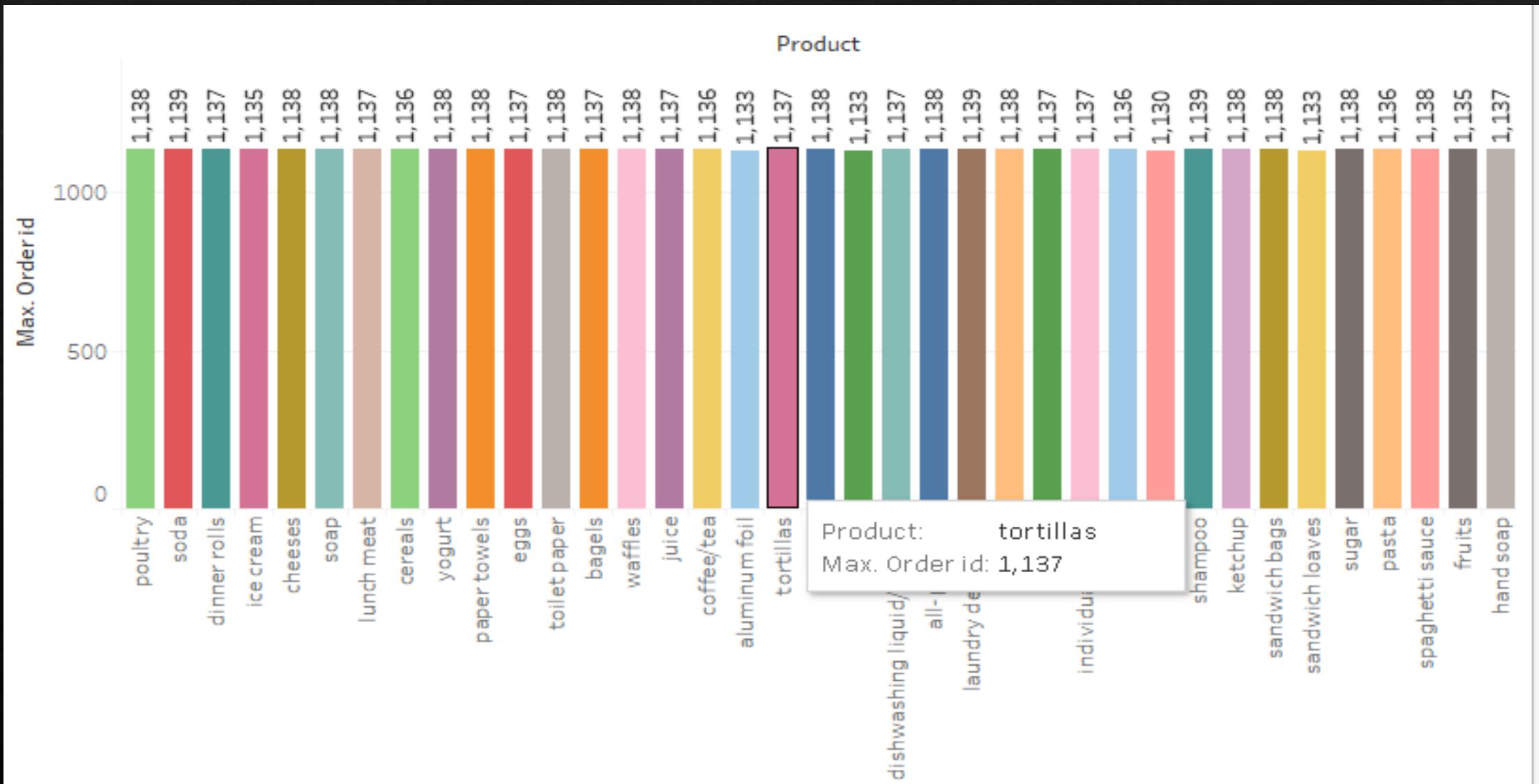


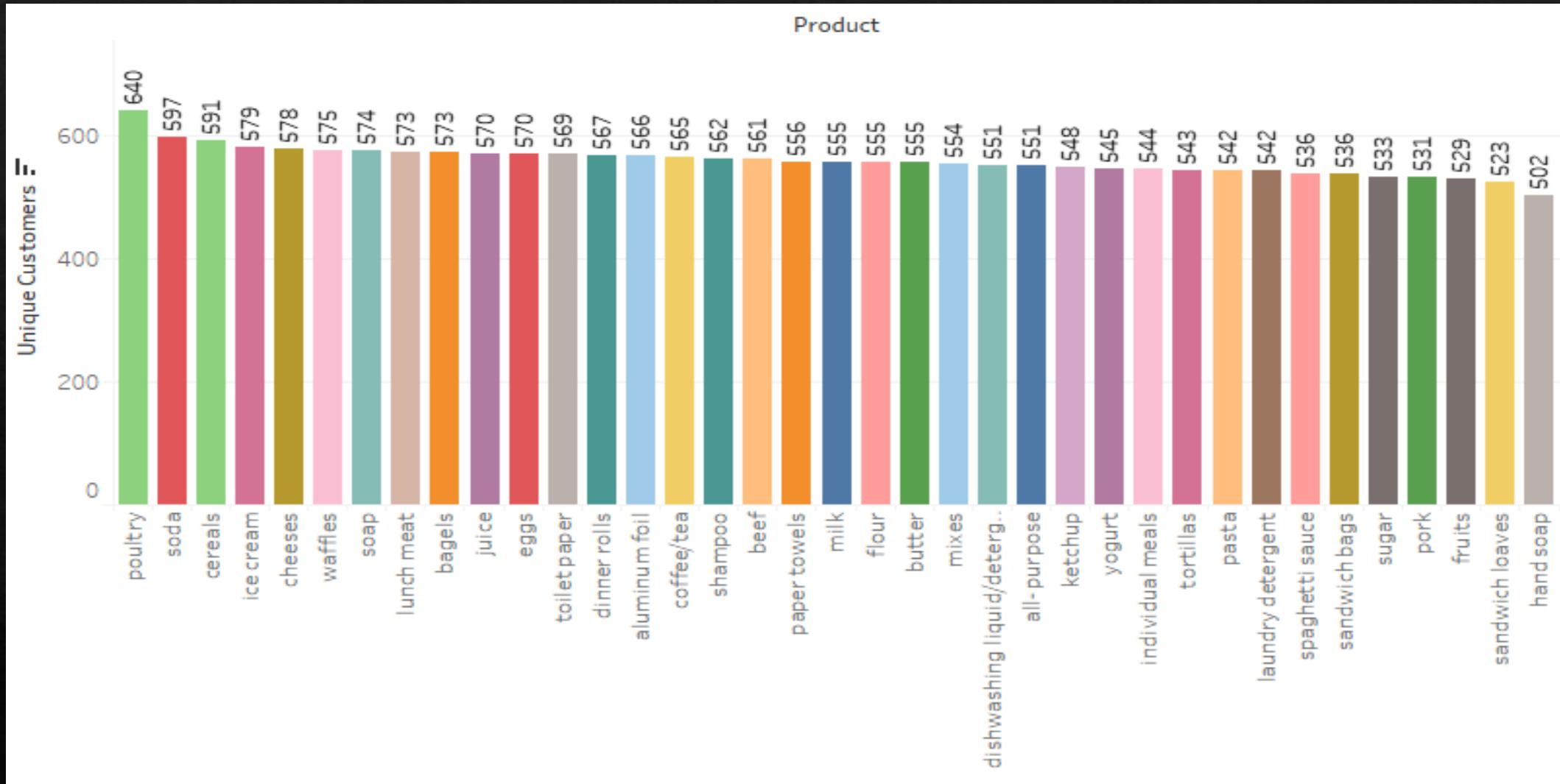
# Sample of the dataset and description of the data

	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
5	2018-01-01	1	flour
6	2018-01-01	1	soda
7	2018-01-01	1	butter
8	2018-01-01	1	beef
9	2018-01-01	1	aluminum foil

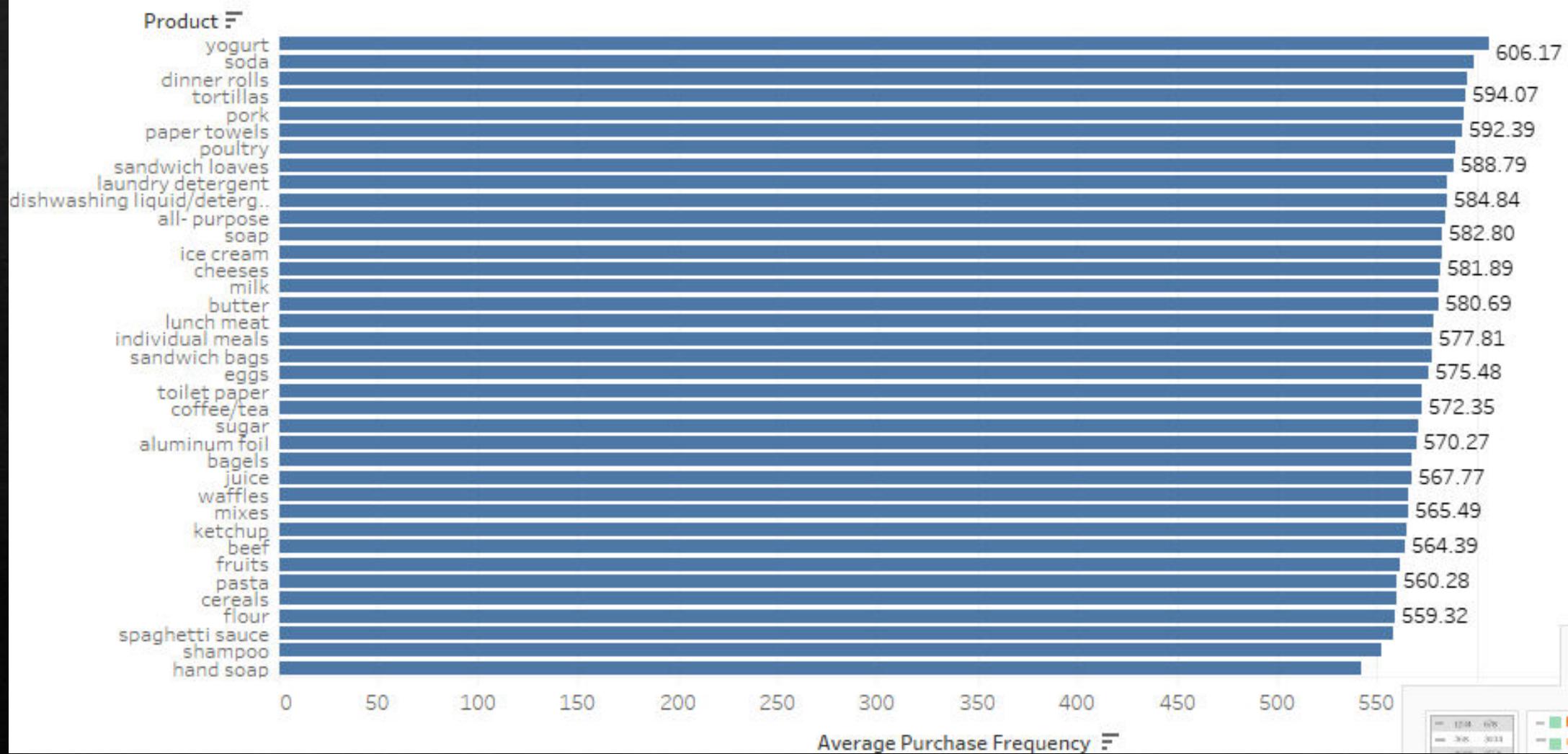
	count	unique	top	freq	first	last	mean	std	min	25%	50%	75%	max
	Date	20641	603	2019-02-08 00:00:00	183	2018-01-01	2020-02-26	NaN	NaN	NaN	NaN	NaN	NaN
	Order_id	20641.0	NaN	NaN	NaN	NaN	NaN	575.986289	328.557078	1.0	292.0	581.0	862.0
	Product	20641	37	poultry	640	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

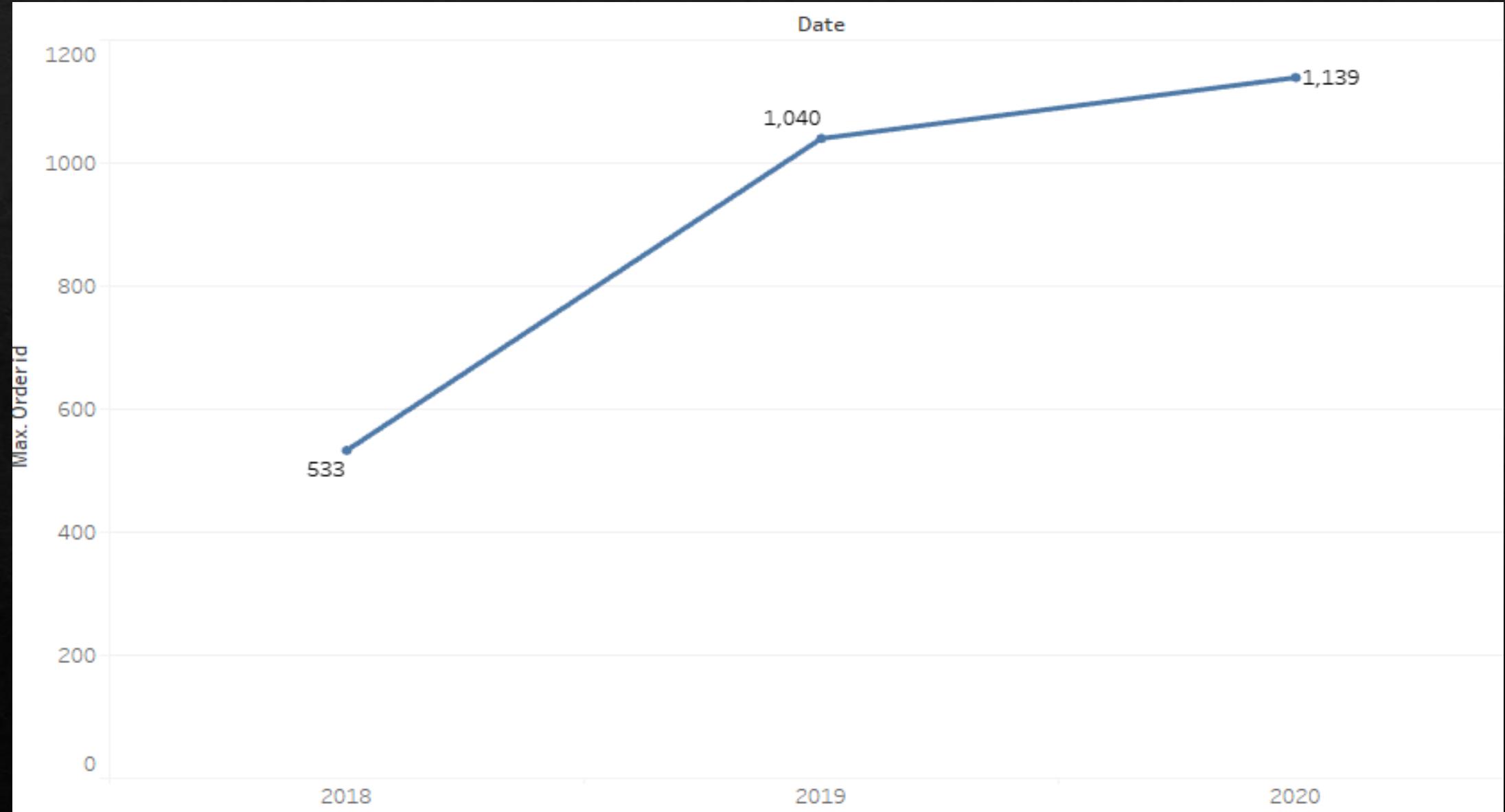
```
# The dataset has 20641 rows and 3 columns 37 unique product categories.  
#The most frequently occurring product category poultry appearing 640 times.  
#It seems that there are missing statistics (mean, std, min, 25%, 50%, 75%, max) for  
the Product column
```



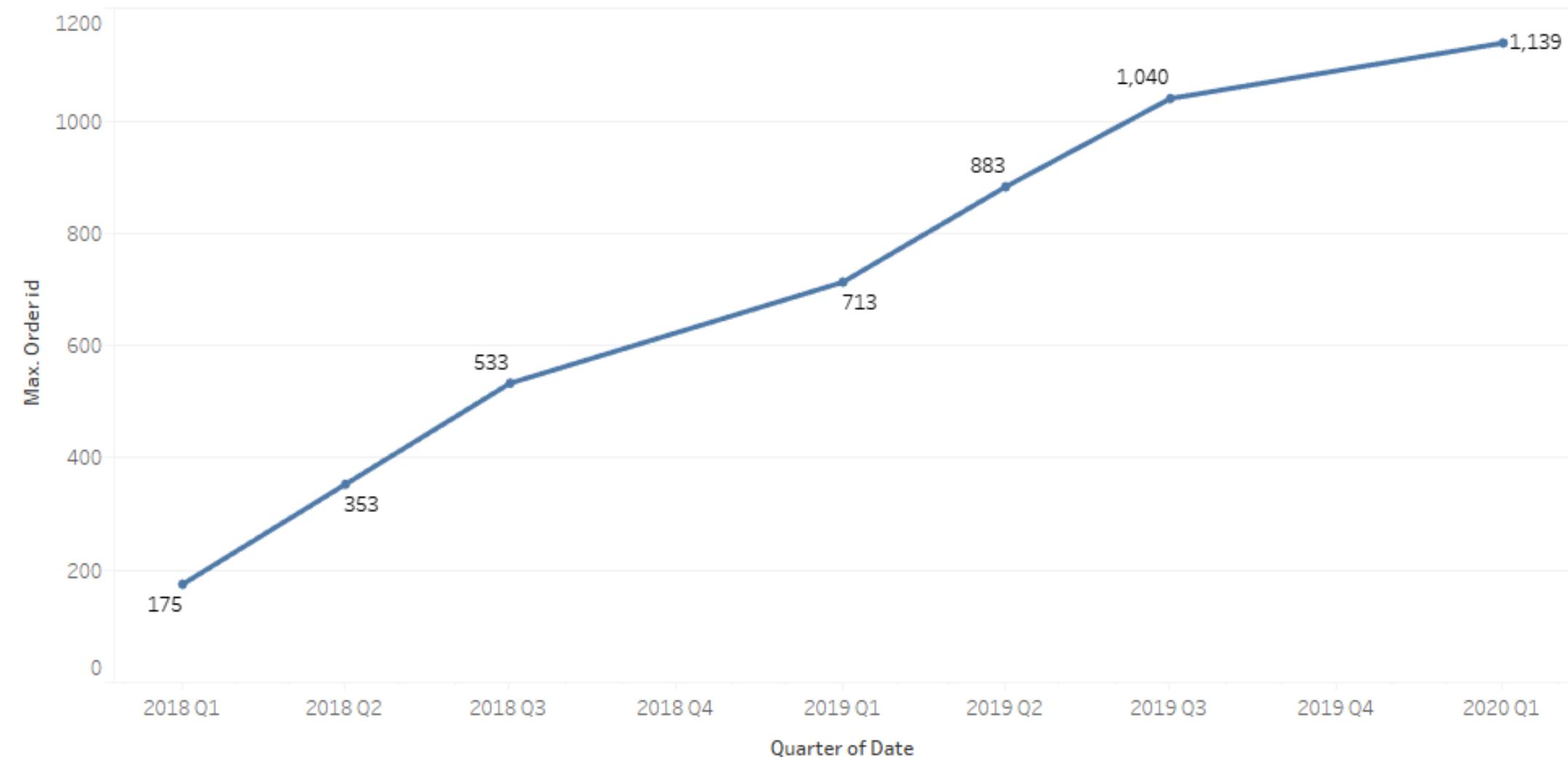


## Avg purchasing Frequency

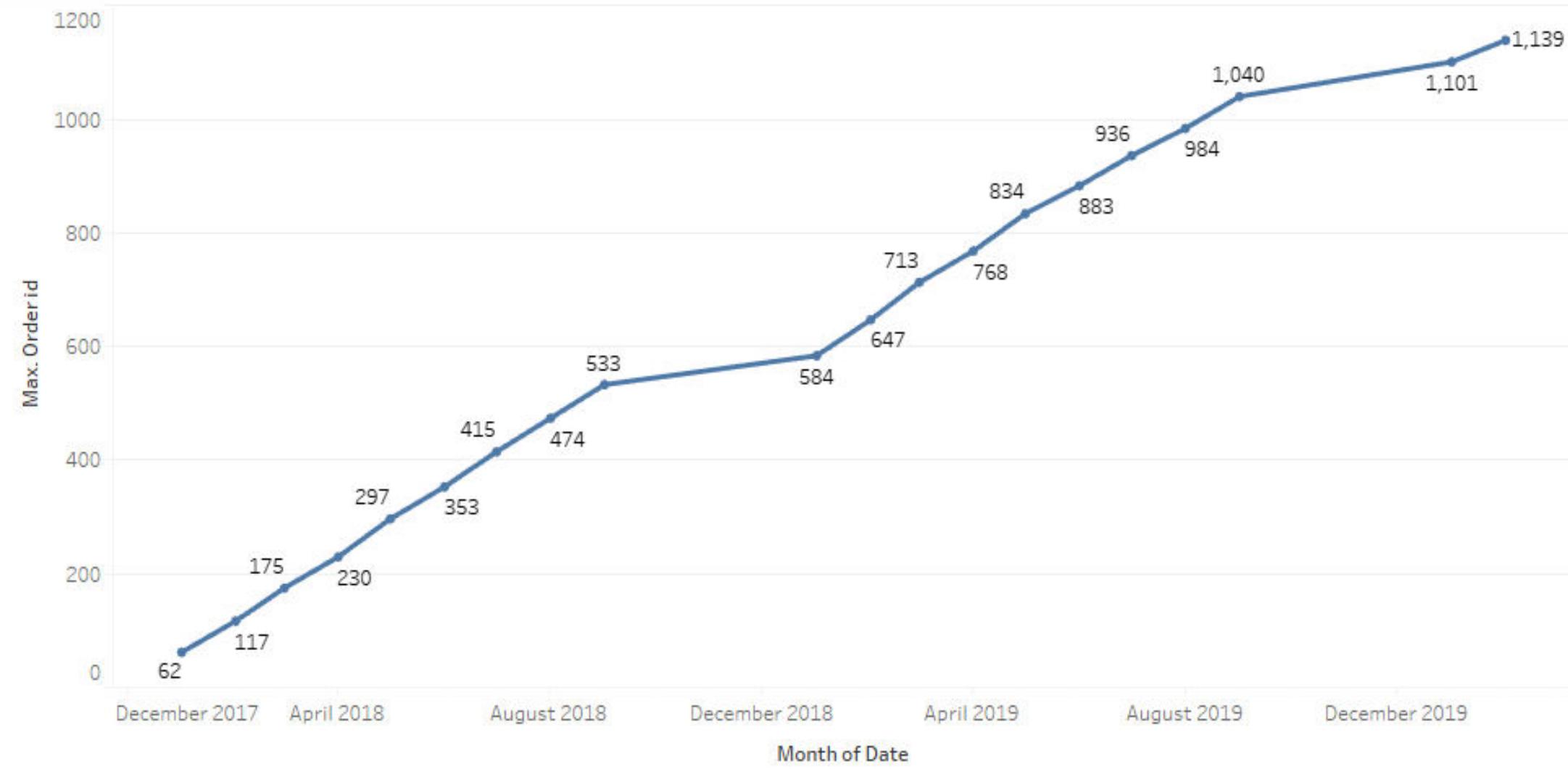




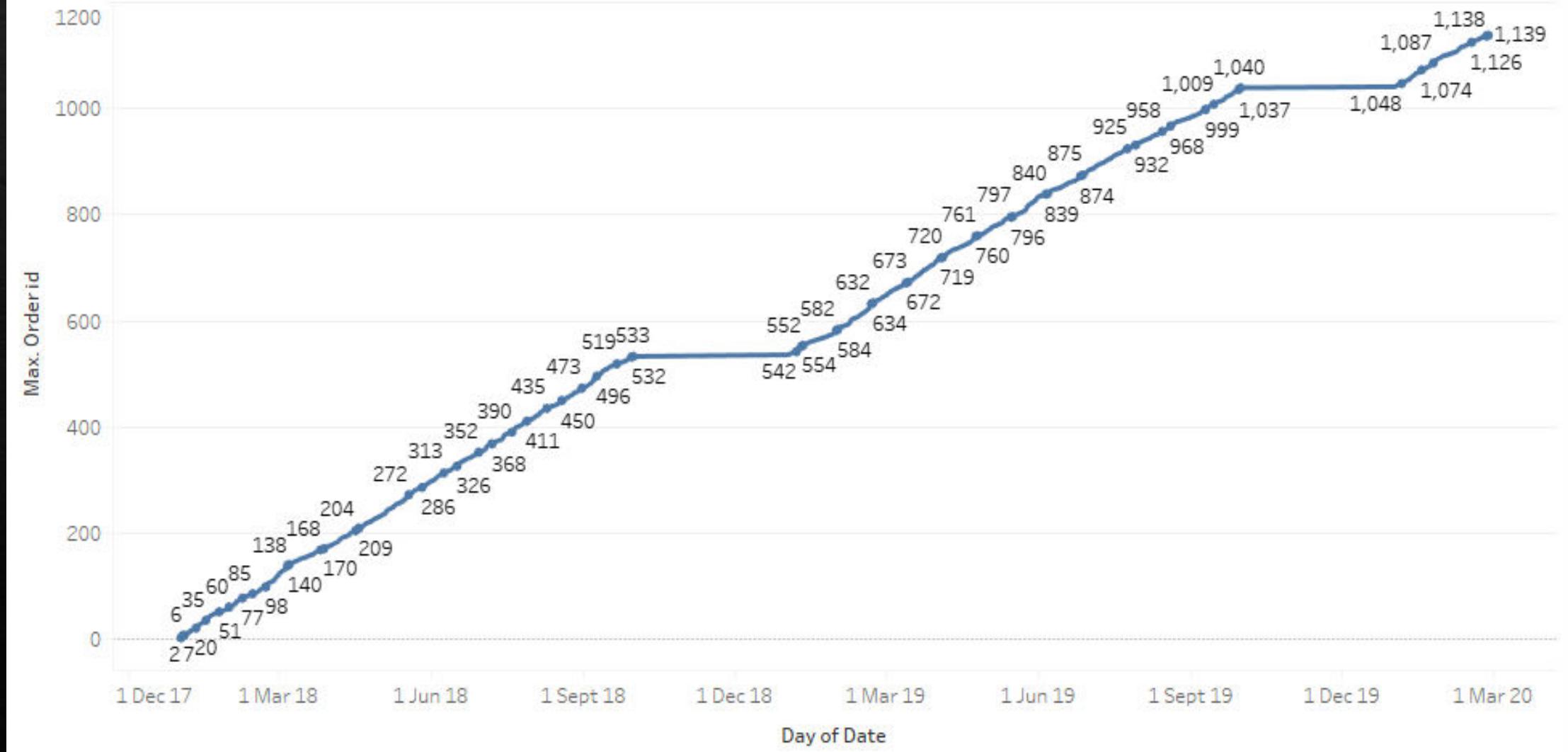
## Quarterly Trend



## Monthly Trend

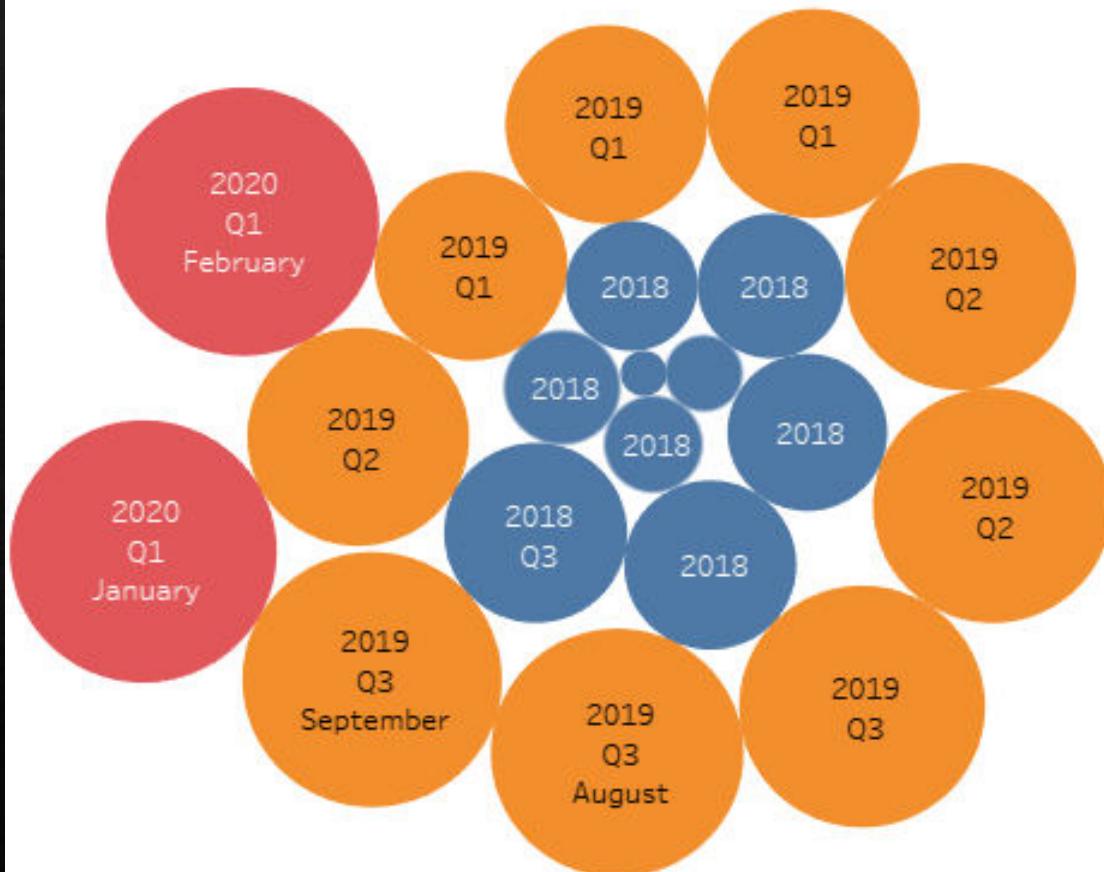


## Day Trend



## Yearly Avg Purchasing FRequency

YEAR(Date)
2018
2019
2020



## Product Avg count



Product
all-purpose
aluminum foil
bagels
beef
butter
cereals
cheeses
coffee/tea
dinner rolls
dishwashing liqui..
eggs
flour
fruits
hand soap
ice cream
individual meals
juice
ketchup
laundry detergent
lunch meat
milk
mixes
paper towels
pasta
green

## Count of Product

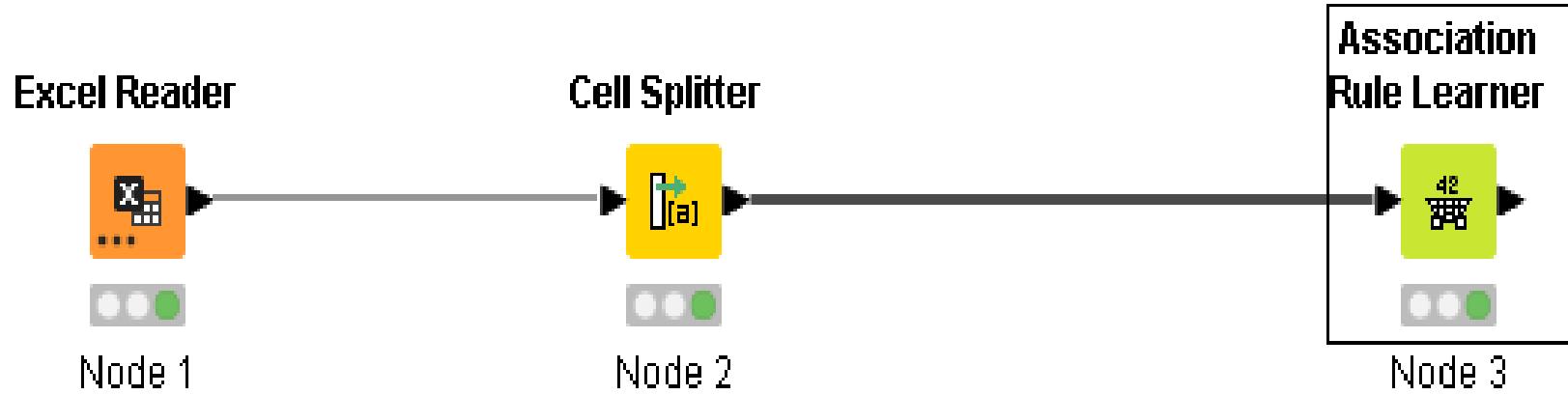
poultry 640	soap 574	dinner rolls 567	butter 555	flour 555	milk 555	mixes 554	all-purpose 551
soda 597	bagels 573	aluminum foil 566		Product: Count of dataset_group (1): butter Count of Order id: 555	butter 555		
cereals 591	lunch meat 573	coffee/tea 565	dishwashing liquid/detergent 551	laundry detergent 542	pasta 542	sandwich bags 536	
ice cream 579	eggs 570	shampoo 562	ketchup 548		spaghetti sauce 536	fruits 529	
cheeses 578	juice 570	beef 561	yogurt 545	sugar 533			
waffles 575	toilet paper 569	papertowels 556	individual meals 544	tortillas 543	pork 531	hand soap 502	

- ❖ Products yogurt, soda, dinner rolls, tortillas, pork, paper-towels, poultry, sandwich loaves, laundry detergent, liquid detergent has the highest frequency(top 10)
- ❖ Data shows raise in orders in consecutive years 2018,2019,2020
- ❖ Quarterly Trend also shows a significant raise in orders
- ❖ Monthly and daily trends has a significant break up of the orders

## Market Basket Analysis



Market basket analysis is a data mining technique used in the field of retail and business analytics to discover patterns, associations, and relationships within transactional data. Its primary objective is to identify which items are frequently purchased together by customers.



Row ID	Date	Order_id	Product	Product_SplitResultSet
Row0	2018-01-01	1	yogurt	[yogurt]
Row1	2018-01-01	1	pork	[pork]
Row2	2018-01-01	1	sandwich bags	[sandwich,bags]
Row3	2018-01-01	1	lunch meat	[lunch,meat]
Row4	2018-01-01	1	all- purpose	[all-,purpose]
Row5	2018-01-01	1	flour	[flour]
Row6	2018-01-01	1	soda	[soda]
Row7	2018-01-01	1	butter	[butter]
Row8	2018-01-01	1	beef	[beef]
Row9	2018-01-01	1	aluminum foil	[aluminum,foil]
Row10	2018-01-01	1	all- purpose	[all-,purpose]
Row11	2018-01-01	1	dinner rolls	[dinner,rolls]
Row12	2018-01-01	1	shampoo	[shampoo]
Row13	2018-01-01	1	all- purpose	[all-,purpose]
Row14	2018-01-01	1	mixes	[mixes]
Row15	2018-01-01	1	soap	[soap]
Row16	2018-01-01	1	laundry detergent	[laundry,detergent]
Row17	2018-01-01	1	ice cream	[ice,cream]
Row18	2018-01-01	1	dinner rolls	[dinner,rolls]
Row19	2018-01-01	2	toilet paper	[toilet,paper]
Row20	2018-01-01	2	shampoo	[shampoo]
Row21	2018-01-01	2	hand soap	[hand,soap]
Row22	2018-01-01	2	waffles	[waffles]
Row23	2018-01-01	2	cheeses	[cheeses]
Row24	2018-01-01	2	mixes	[mixes]
Row25	2018-01-01	2	milk	[milk]
Row26	2018-01-01	2	sandwich bags	[sandwich,bags]
Row27	2018-01-01	2	laundry detergent	[laundry,detergent]
Row28	2018-01-01	2	dishwashing liquid/dete...	[dishwashing,liquid/detergent]
Row29	2018-01-01	2	waffles	[waffles]
Row30	2018-01-01	2	individual meals	[individual,meals]
Row31	2018-01-01	2	hand soap	[hand,soap]
Row32	2018-01-01	2	individual meals	[individual,meals]
Row33	2018-01-01	2	yogurt	[yogurt]

## Identified Associations:

Row ID	Support	Confidence	Lift	Consequence	Implies	Items
rule3	0.026	1	38.509	spaghetti	<---	[sauce]
rule4	0.026	1	38.509	sauce	<---	[spaghetti]
rule5	0.026	1	38.083	laundry	<---	[detergent]
rule6	0.026	1	38.083	detergent	<---	[laundry]
rule7	0.026	1	37.943	individual	<---	[meals]
rule8	0.026	1	37.943	meals	<---	[individual]
rule9	0.027	1	37.461	all-	<---	[purpose]
rule10	0.027	1	37.461	purpose	<---	[all-]
rule11	0.027	1	37.461	dishwashing	<---	[liquid/deter...]
rule12	0.027	1	37.461	liquid/deterg...	<---	[dishwashing]
rule14	0.027	1	36.468	aluminum	<---	[foil]
rule15	0.027	1	36.468	foil	<---	[aluminum]
rule16	0.027	1	36.404	dinner	<---	[rolls]
rule17	0.027	1	36.404	rolls	<---	[dinner]
rule19	0.028	1	36.023	lunch	<---	[meat]
rule20	0.028	1	36.023	meat	<---	[lunch]
rule21	0.028	1	35.649	ice	<---	[cream]
rule22	0.028	1	35.649	cream	<---	[ice]
rule1	0.025	1	19.491	sandwich	<---	[loaves]
rule2	0.026	1	19.491	sandwich	<---	[bags]
rule0	0.024	1	19.183	soap	<---	[hand]
rule13	0.027	1	18.348	paper	<---	[towels]
rule18	0.028	1	18.348	paper	<---	[toilet]

Row ID	D Support	D Confide...	D ▼ Lift	S Conseq...	S implies	[...] Items
rule3	0.026	1	38.509	spaghetti	<-->	[sauce]
rule4	0.026	1	38.509	sauce	<-->	[spaghetti]
rule5	0.026	1	38.083	laundry	<-->	[detergent]
rule6	0.026	1	38.083	detergent	<-->	[laundry]
rule7	0.026	1	37.943	individual	<-->	[meals]
rule8	0.026	1	37.943	meals	<-->	[individual]
rule9	0.027	1	37.461	all-	<-->	[purpose]
rule10	0.027	1	37.461	purpose	<-->	[all-]
rule11	0.027	1	37.461	dishwashing	<-->	[liquid/deter...]
rule12	0.027	1	37.461	liquid/deterg...	<-->	[dishwashing]
rule14	0.027	1	36.468	aluminum	<-->	[foil]
rule15	0.027	1	36.468	foil	<-->	[aluminum]
rule16	0.027	1	36.404	dinner	<-->	[rolls]
rule17	0.027	1	36.404	rolls	<-->	[dinner]
rule19	0.028	1	36.023	lunch	<-->	[meat]
rule20	0.028	1	36.023	meat	<-->	[lunch]
rule21	0.028	1	35.649	ice	<-->	[cream]
rule22	0.028	1	35.649	cream	<-->	[ice]
rule1	0.025	1	19.491	sandwich	<-->	[loaves]
rule2	0.026	1	19.491	sandwich	<-->	[bags]
rule0	0.024	1	19.183	soap	<-->	[hand]
rule13	0.027	1	18.348	paper	<-->	[towels]
rule18	0.028	1	18.348	paper	<-->	[toilet]

- ❖ In total of 22 rules found in the dataset , A lift value greater than 1 suggests a positive association. A confidence of 1.0 means that the antecedent is always bought when the consequent is purchased. Support measures the frequency of occurrence of both items together. It is a proportion of transactions that contain both items.
- ❖ The items on both sides of the "<-->" arrow are the antecedent and consequent of the association rule, indicating which items tend to be purchased together.

- ❖ Based on these association rules here are some potential recommendations based on the provided rules:
  - ❖ Spaghetti & Sauce: Since spaghetti and sauce have a high lift and confidence, you can recommend placing them together in-store or running promotions like "Buy Spaghetti, Get Sauce at a Discount."
  - ❖ Laundry & Detergent: Given the strong association between laundry and detergent, you could consider bundling these products or offering detergent discounts to laundry product purchasers.
  - ❖ Individual Meals: If individual meals and meals are frequently bought together, you might create meal combos for customers looking for quick meal solutions.
  - ❖ All-Purpose & Dishwashing Liquid/Detergent: Encourage customers to purchase both all-purpose and liquid dishwashing detergent together, perhaps by offering a discount on the combination.
  - ❖ Aluminum Foil: Since aluminum and foil have a strong association, consider placing them together in-store or running promotions on foil for customers buying aluminum products.

- ❖ Dinner Rolls: Promote dinner rolls alongside dinner items, especially during dinner hours or special meal occasions.
- ❖ Lunch & Meat: If lunch and meat are frequently bought together, create lunch kits or promotions that include meat options.
- ❖ Ice Cream: Since ice and cream have a high lift, you could promote ice cream during the purchase of ice products, especially during warm weather.
- ❖ Sandwich & Bags: Encourage the purchase of sandwich bags when customers buy sandwich ingredients.
- ❖ Soap & Hand: Place hand soap and soap products together in-store for convenience.
- ❖ Paper Towels & Toilet Paper: Consider bundling paper towel and toilet paper products or running promotions for customers buying one when purchasing the other.
- ❖ These recommendations are based on the associations observed in the dataset and can be used to optimize product placement, marketing campaigns, and pricing strategies to increase sales and customer satisfaction.

Thank  
you