



MARKET BASKET ANALYSIS

Zinia Gaba



AGENDA

- Exploratory Data Analysis –Product Mix, Trends over time
- Association Rules, Threshold parameters, Workflow
- Recommendations



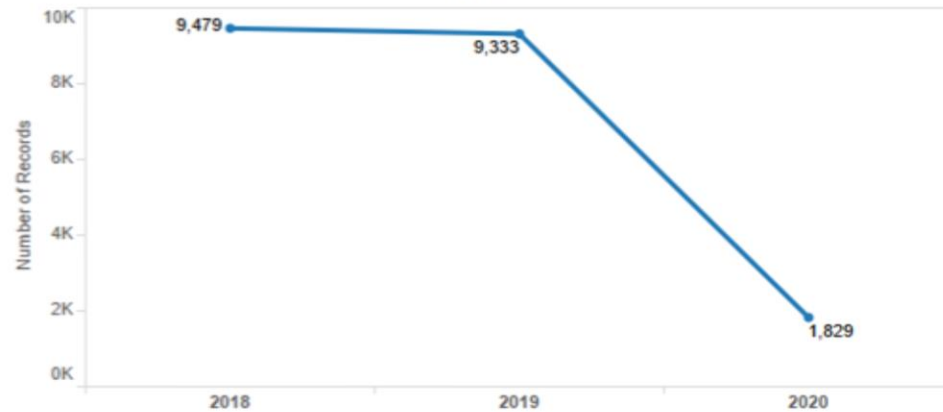
PRODUCT MIX

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					
cereals 591	lunch meat 573	coffee/tea 565	dishwashing liquid/ 551	tortillas 543	pasta 542	laundry det 542	sandwich b 536
ice cream 579	eggs 570	shampoo 562	ketchup 548	spaghetti sauce 536		fruits 529	sandwich l 523
cheeses 578	juice 570	beef 561	yogurt 545	sugar 533			
waffles 575	toilet paper 569	paper towels 556	individual meals 544	pork 531		hand soap 502	

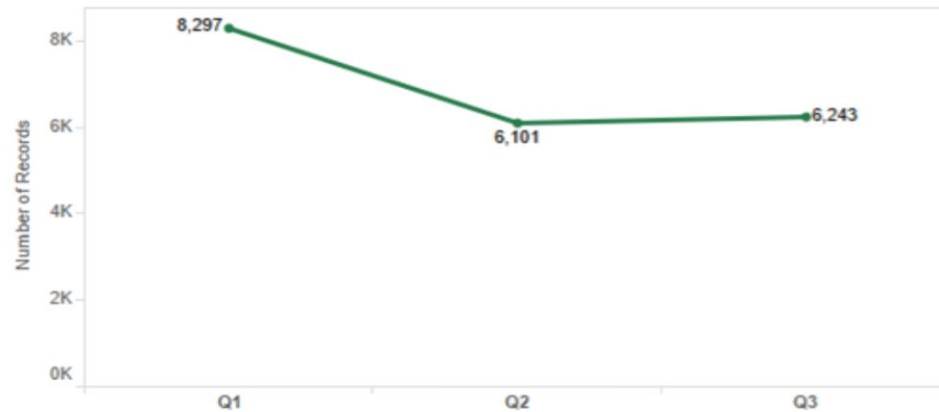


TRENDS OVER TIME

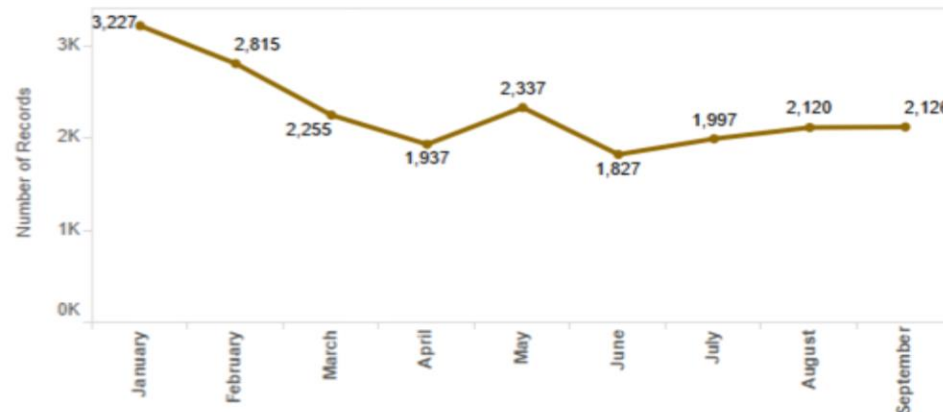
Year



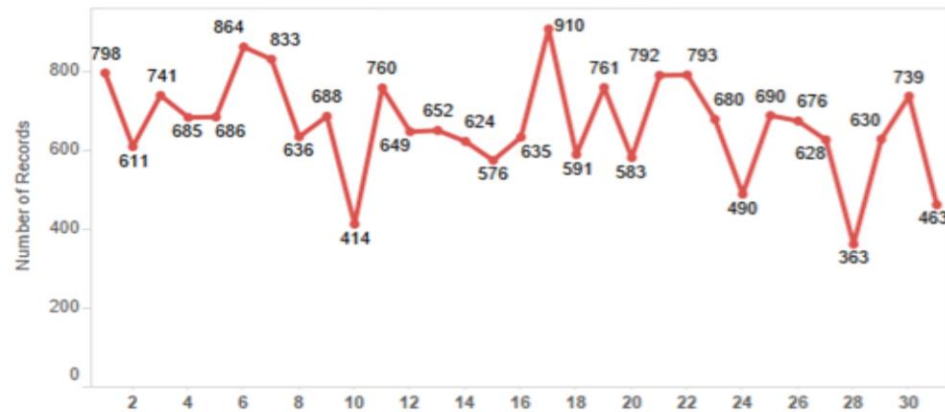
Quarter



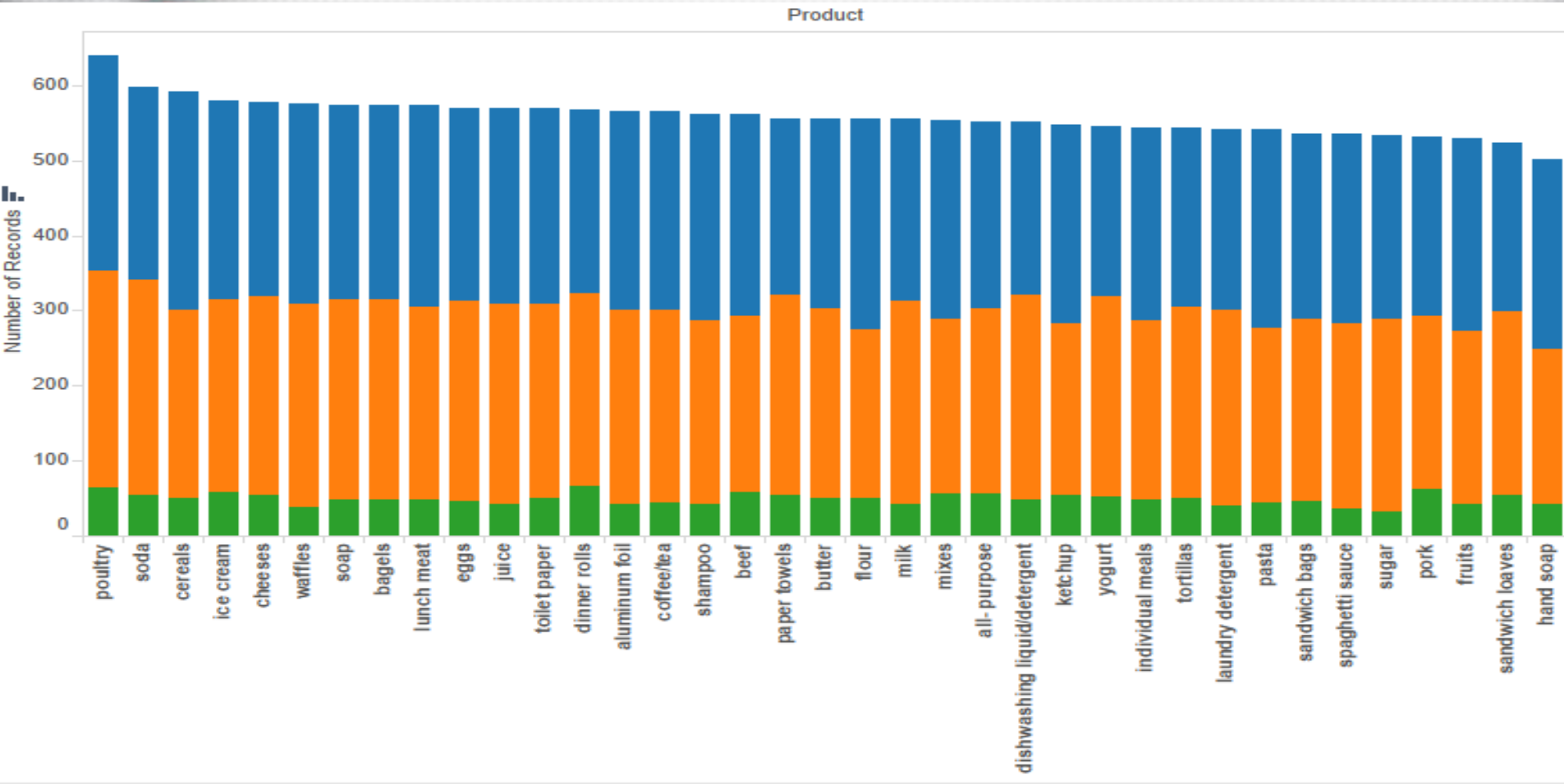
Month



Date



EDA- Product Trends Over Years





INSIGHTS

1. Highest number of products bought in 2018, with a slight decline in 2019 and heavy decline in 2020
2. Quarter 1 has the highest number of product sold, followed by Quarter 2 and Quarter 3(No records for Q4)
3. January has highest number of records and June the lowest
4. Product brought maximum times – ‘Poultry’, least times – ‘hand soap’
5. There seems to be no unusual trends of products over the years. All products follow the total yearly trend



Association Rules and MBA Recommendations



ASSOCIATION PARAMETERS

Association Rules

It shows the frequently brought items in our retail data set.

Support

This metric gives the popularity of the set A across the all transactions

Confidence

This metric gives the likelihood of a customer purchasing both set A and set B together in a transaction

Lift

The increase in the probability of a customer buying set B provided they have bought set A



THRESHOLD VALUES

- **Support** – We would prefer a higher support as higher popularity would mean we can apply these rules to multiple transactions and our association rules are scalable
- **Confidence** – We would also prefer a higher confidence value as it greater number of appearance of the combination higher the relevance of the rules
- **Lift** – A higher lift is absolutely necessary(>1) as it establishes an increase or lift in the probability.

Threshold Selected

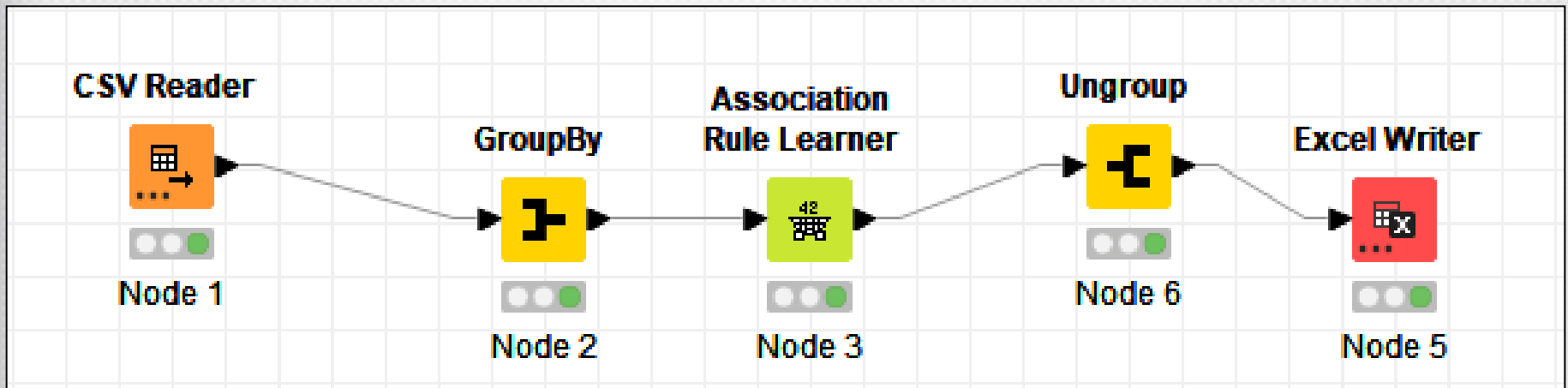
Confidence-60%

Support-5%

Lift- >1



KNIME WORKFLOW



Used **Knime** for Market Basket Analysis. Process followed :

- Read the file
- Grouped the data by basket , i.e. Order ID and created a set of products
- Association rule with support = 0.005 and confidence = 0.60
- Ungrouped the data and exported



ASSOCIATIONS

Basket	Support	Confidence	Lift	Consequent	Items
20	0.050921861	0.617021277	1.55828655	cereals	cheeses,bagels,sandwich bags
3	0.050921861	0.674418605	1.726208518	cheeses	bagels,cereals,sandwich bags
15	0.050043898	0.612903226	1.615964755	coffee/tea	yogurt,cheeses,cereals
7	0.053555751	0.642105263	1.650920756	dinner rolls	spaghetti sauce,poultry,laundry detergent
8	0.051799824	0.641304348	1.648861517	dinner rolls	spaghetti sauce,poultry,ice cream

Association table with top 5 rows(alphabetical). Inference of row-1, basket -20:

- When items – Cheeses , bagels and sandwich bags are bought together , the probability of buying cereals gets increases 1.55 times.
- The percentage of transactions which contain Cheeses, bagels and sandwich bags is 5.09% (Support)
- Cereal is transacted 61.7% of the times when Cheeses, bagels and sandwich bags are bought (Confidence)



TOP 10 BUNDLING

#	Antecedent	Consequent
1	eggs, ice cream, pasta	paper towels
2	paper towels, eggs, ice cream	pasta
3	bagels, cereals, sandwich bags	cheeses
4	yogurt, toilet paper, aluminum foil	juice
5	yogurt, poultry, aluminum foil	mixes
6	cheeses, bagels, cereals	sandwich bags
7	spaghetti sauce, poultry, laundry detergent	dinner rolls
8	spaghetti sauce, poultry, ice cream	dinner rolls
9	yogurt, poultry, aluminum foil	juice
10	dinner rolls, spaghetti sauce, ice cream	poultry

Offer the above Consequent item whenever a customer does a transaction of these Antecedent items



RECOMM- CONTD

- Offer discounts and offers at the end of the month, i.e. 28th onwards and the month of June
- Need to increase sales of 'Hand soaps'. The likelihood of customer buying hand soaps will increase when bought with following products. (Support = 6%, Confidence = 40%, lift >1)

[pasta, pork]	hand soap
[ice cream, sandwich loaves]	hand soap
[toilet paper, ketchup]	hand soap
[cheeses, pasta]	hand soap
[milk, ketchup]	hand soap
[spaghetti sauce, individual meals]	hand soap
[pasta, tortillas]	hand soap
[bagels, pasta]	hand soap
[pasta, coffee/tea]	hand soap