Marketing & Retail Analytics

Part B - MBA

Exploratory Analysis

Problem Statement

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.

Summary (Info, Shape, Statistical Summary)

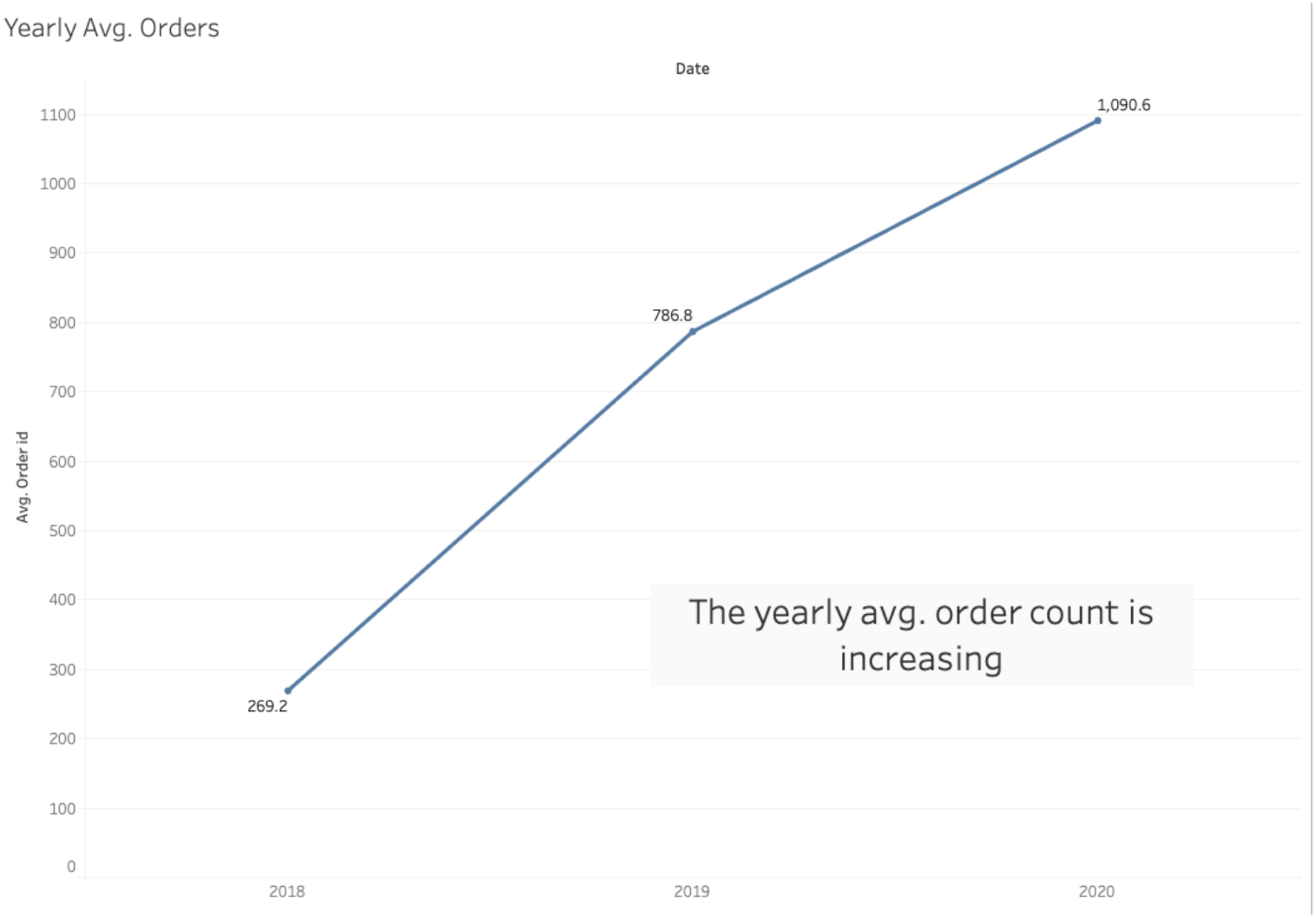
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20641 entries, 0 to 20640
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   Date      20641 non-null  object
 1   Order_id  20641 non-null  int64 
 2   Product   20641 non-null  object
dtypes: int64(1), object(2)
memory usage: 483.9+ KB
```

```
The numner of rows in the data set =  20641
The numner of columns in the data set =  3
```

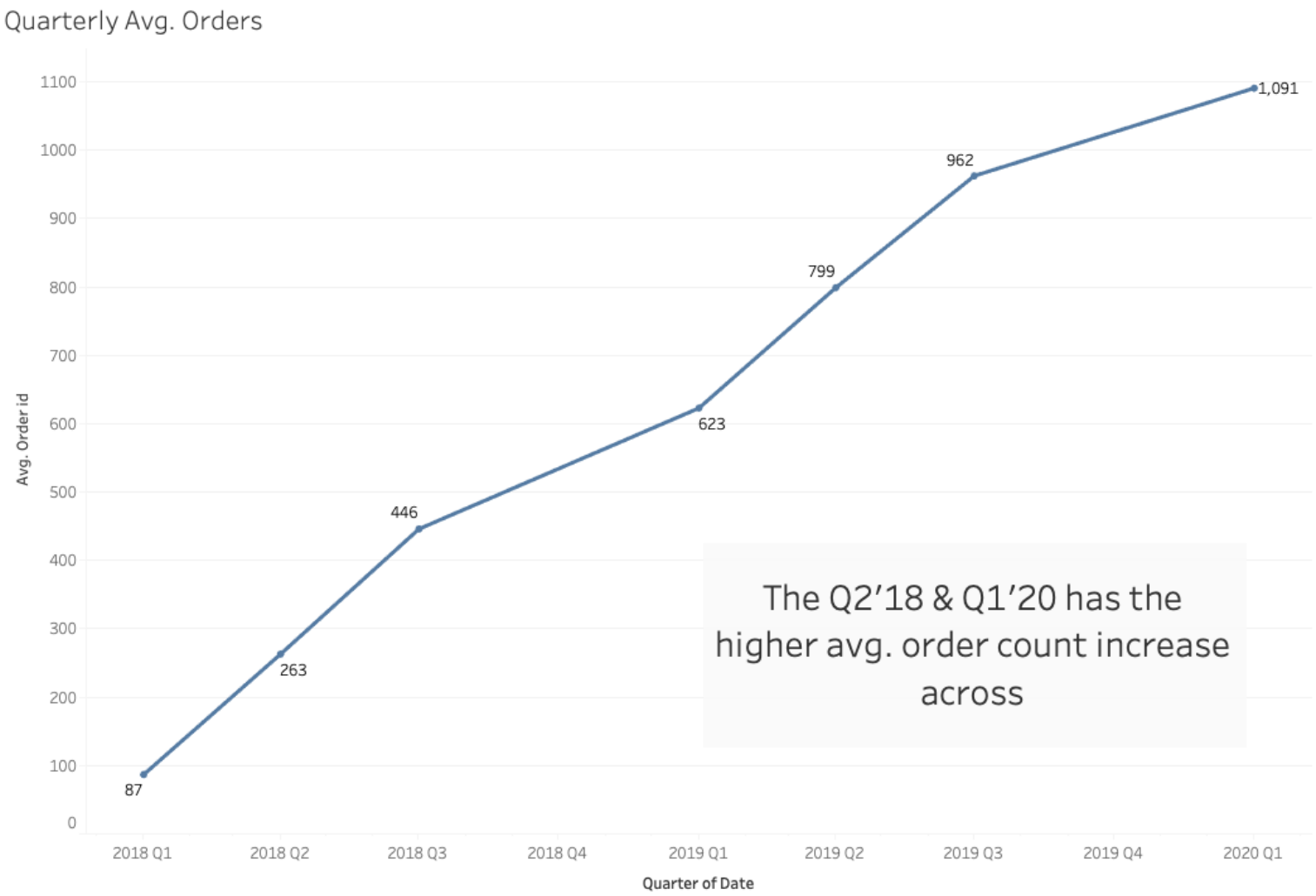
	count	mean	std	min	25%	50%	75%	max
Order_id	20641.0	575.99	328.56	1.0	292.0	581.0	862.0	1139.0

	count	unique	top	freq
Date	20641	603	08-02-2019	183
Product	20641	37	poultry	640

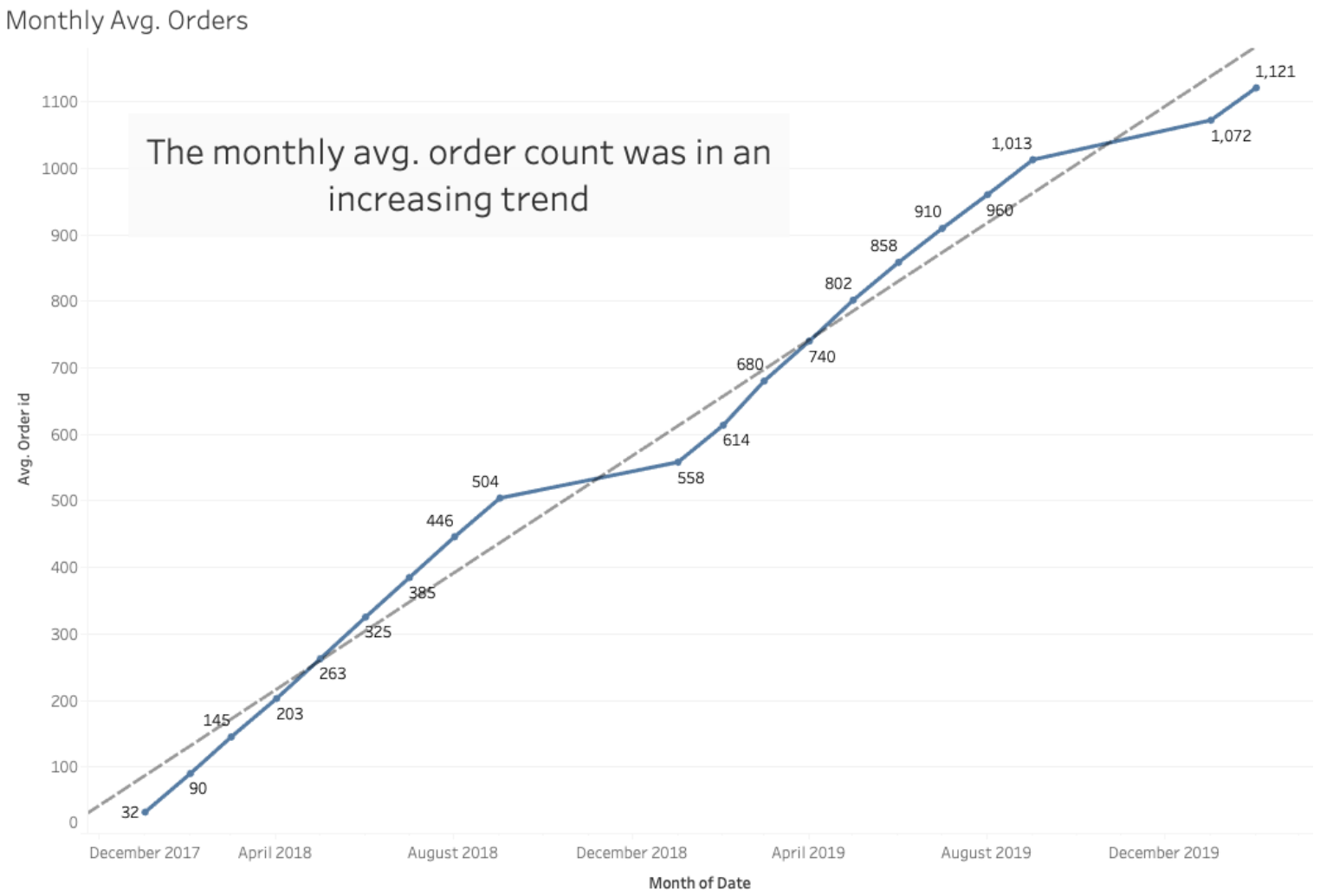
Yearly Trend



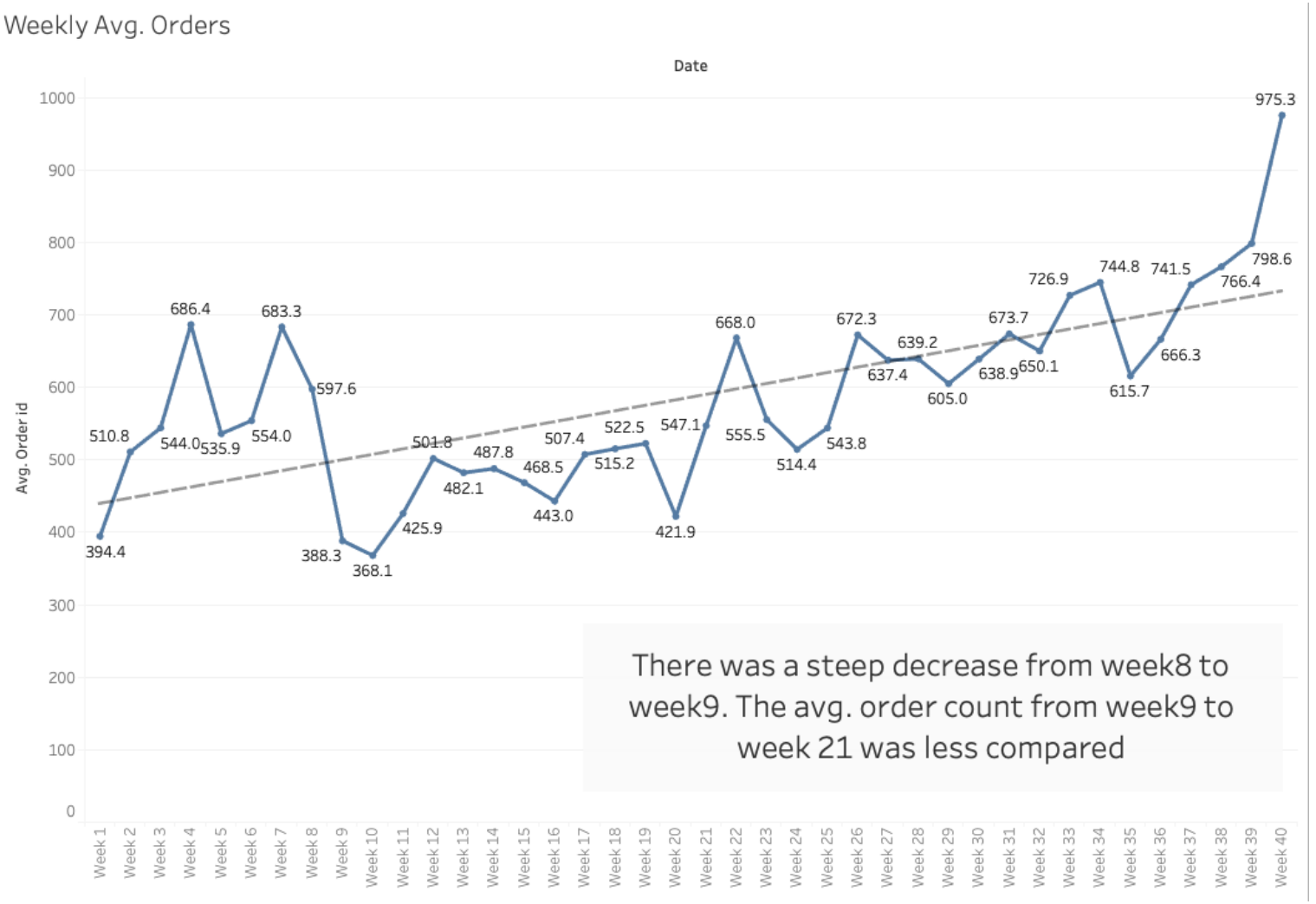
Quarterly Trend



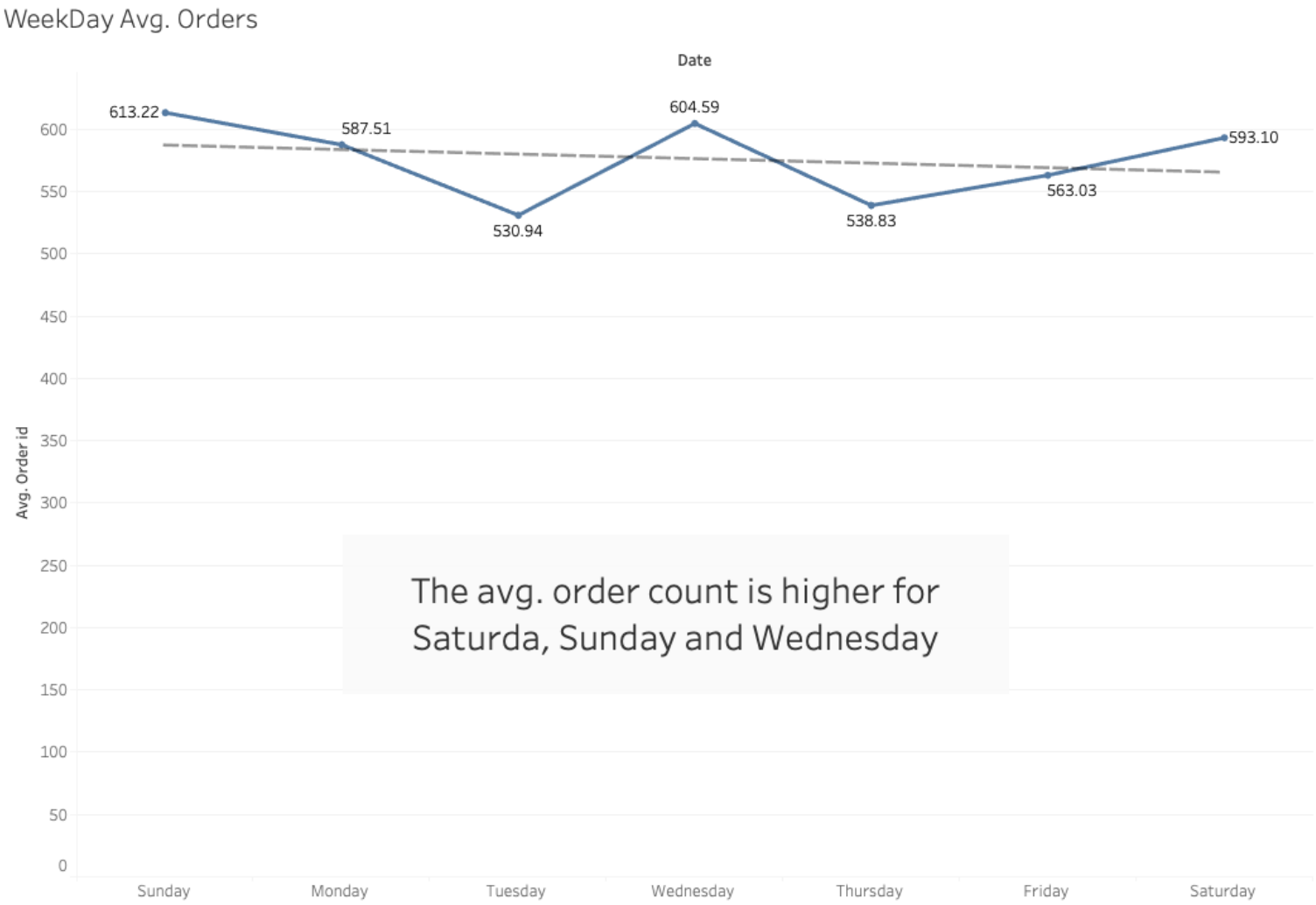
Monthly Trend



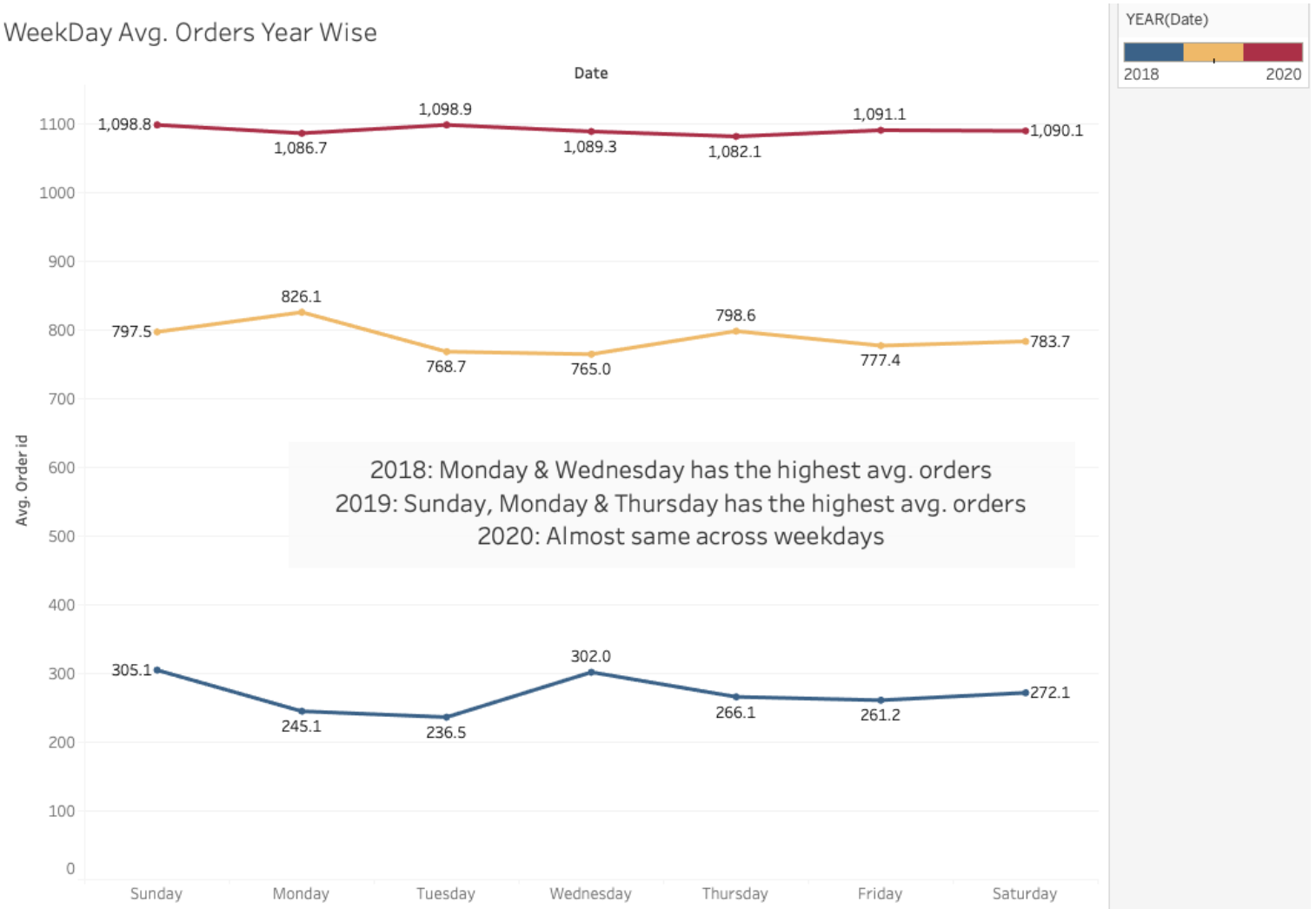
Weekly Trend



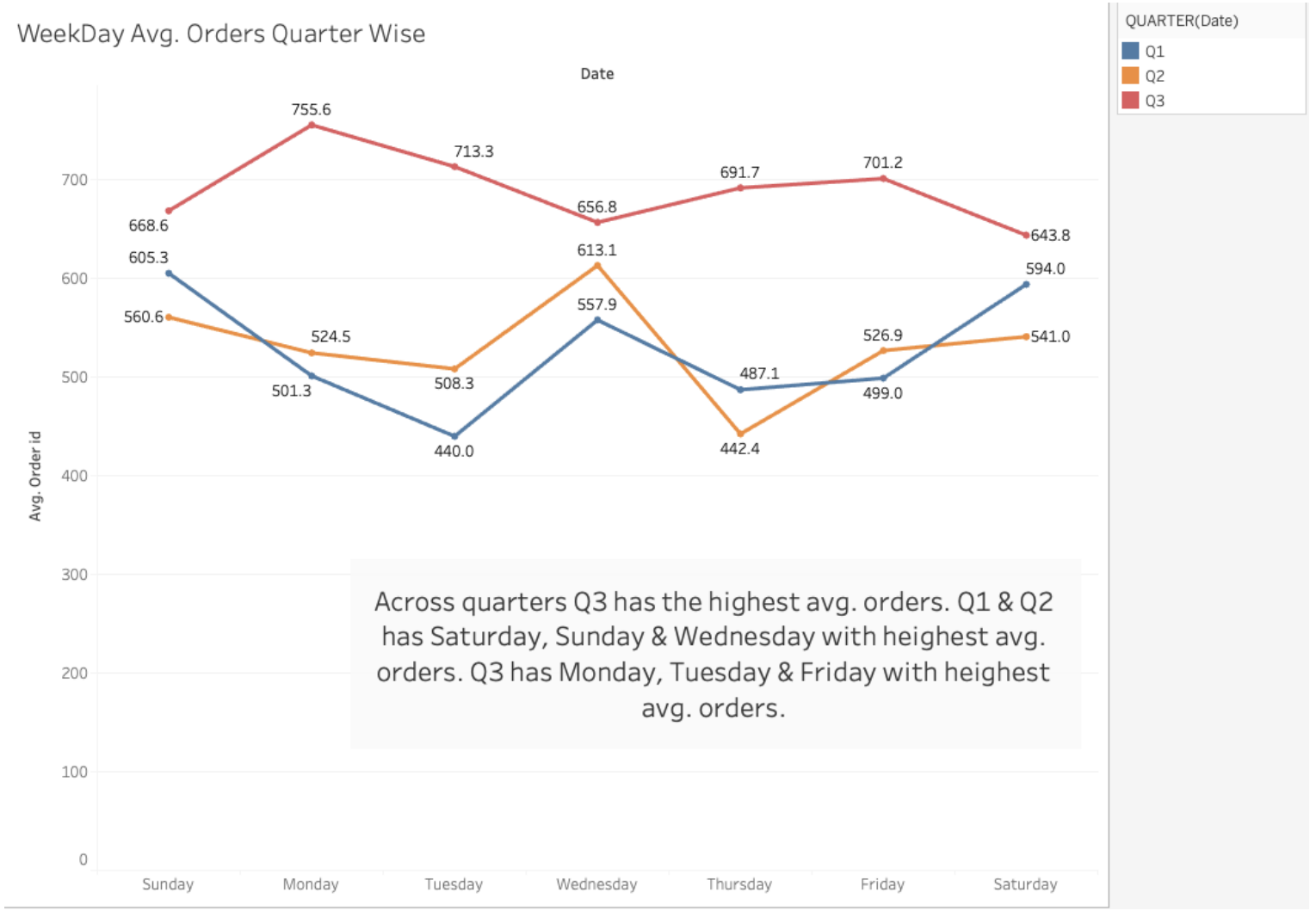
Weekday Trend



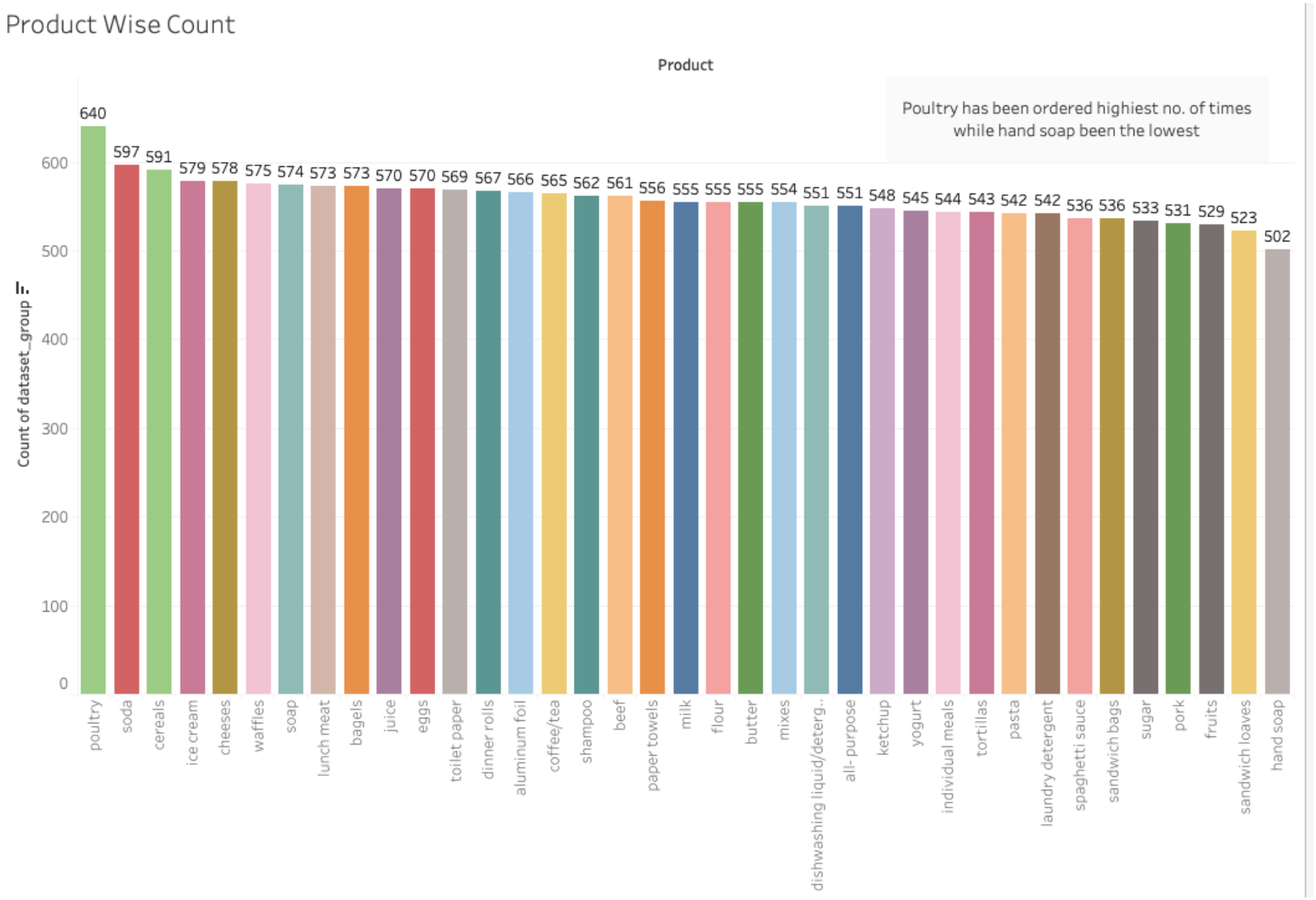
Weekday Yearly Trend



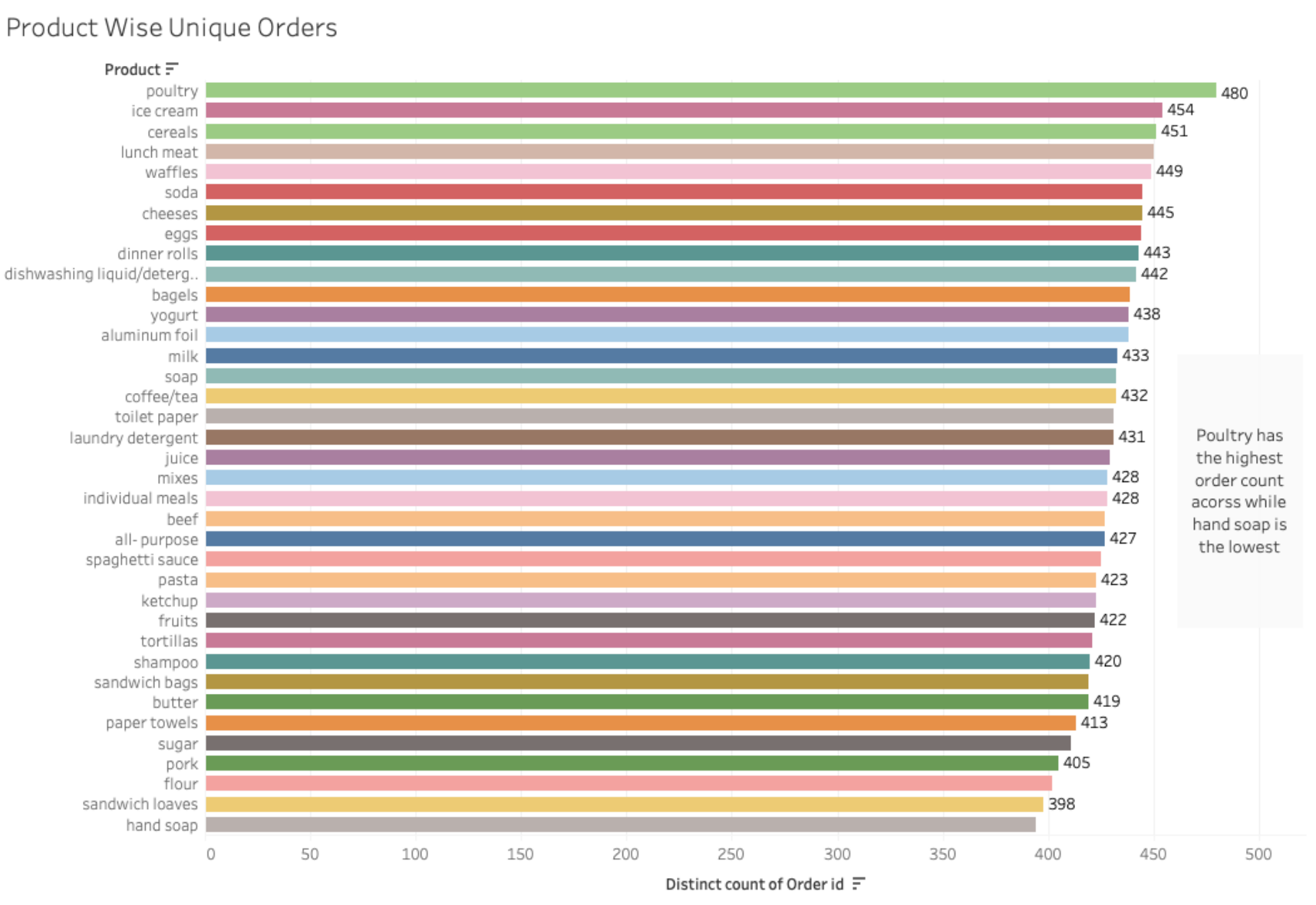
Weekday Quarterly Trend



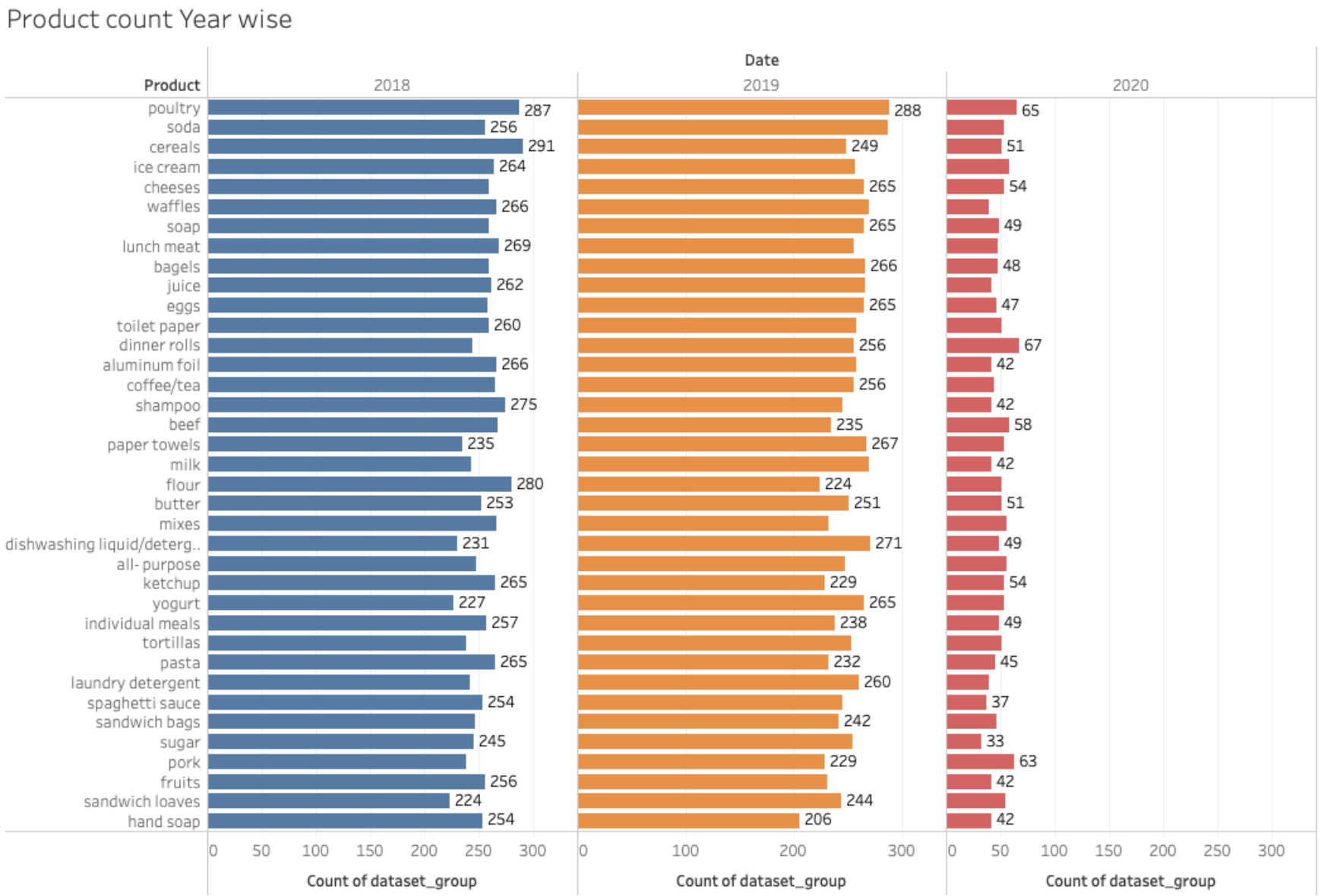
Product Wise Count



Product Wise Orders



Product Wise Orders



Product Count Month Wise

Product Count Month Wise

Product	Date								
	January	February	March	April	May	June	July	August	September
all- purpose	91	77	58	55	62	46	48	53	61
aluminum foil	85	70	59	62	53	56	55	59	67
bagels	94	80	73	60	56	56	40	64	50
beef	87	89	67	43	64	58	45	57	51
butter	86	64	73	43	83	48	42	49	67
cereals	95	82	57	52	72	49	52	74	58
cheeses	91	87	67	50	44	65	56	58	60
coffee/tea	94	60	60	50	59	64	68	52	58
dinner rolls	94	81	60	67	60	41	65	53	46
dishwashing liquid/deter..	90	87	69	52	60	41	49	51	52
eggs	86	67	78	59	70	47	49	51	63
flour	89	65	57	55	62	49	69	49	60
fruits	72	75	55	48	70	50	44	72	43
hand soap	82	81	50	44	56	40	50	46	53
ice cream	95	83	59	61	47	53	60	67	54
individual meals	85	65	61	58	51	47	46	58	73
juice	91	66	74	60	70	36	48	71	54
ketchup	91	80	59	43	70	42	52	51	60
laundry detergent	82	67	60	41	61	57	43	69	62
lunch meat	78	86	51	52	57	51	79	59	60
milk	84	81	53	54	67	48	57	56	55
mixes	90	93	59	42	52	49	57	55	57
paper towels	92	77	60	54	75	46	53	41	58
pasta	83	67	50	66	65	51	60	47	53
pork	92	77	60	44	48	44	54	50	62
poultry	94	94	70	65	66	55	63	60	73
sandwich bags	67	82	66	46	67	51	48	53	56
sandwich loaves	99	69	51	48	61	45	58	51	41
shampoo	97	74	64	44	51	55	61	60	56
soap	85	70	70	55	68	41	55	69	61
soda	97	79	56	55	72	48	52	69	69
spaghetti sauce	73	65	73	60	61	47	52	57	48
sugar	75	79	54	40	56	53	59	51	66
toilet paper	86	79	63	54	78	54	49	51	55
tortillas	85	65	53	47	76	52	56	55	54
waffles	91	82	55	48	71	46	53	73	56
yogurt	79	70	51	60	76	46	50	59	54

CNT(datase...
36 99

Product Count Weekday Wise

Product Count WeekDay Wise

Product	Date						
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
all- purpose	87	82	81	75	73	74	79
aluminum foil	91	69	84	68	96	86	72
bagels	90	69	95	67	93	88	71
beef	87	67	73	73	81	83	97
butter	77	79	63	80	89	71	96
cereals	96	73	84	73	97	93	75
cheeses	84	67	79	81	93	95	79
coffee/tea	87	59	98	94	85	75	67
dinner rolls	69	71	91	82	82	88	84
dishwashing liquid/deter..	80	67	70	83	75	89	87
eggs	86	66	83	91	82	86	76
flour	87	75	77	77	73	83	83
fruits	78	65	64	83	77	82	80
hand soap	67	60	63	78	73	85	76
ice cream	87	81	82	81	89	87	72
individual meals	86	73	79	81	73	84	68
juice	97	74	79	80	75	87	78
ketchup	87	58	77	91	73	79	83
laundry detergent	87	75	66	89	72	70	83
lunch meat	98	69	87	66	85	80	88
milk	84	67	90	75	80	89	70
mixes	84	67	73	78	89	88	75
paper towels	80	70	80	68	82	83	93
pasta	70	68	74	95	71	72	92
pork	89	68	83	65	74	79	73
poultry	82	99	82	82	88	102	105
sandwich bags	75	75	76	66	70	84	90
sandwich loaves	81	68	89	76	71	78	60
shampoo	88	67	89	91	72	75	80
soap	89	76	81	93	72	74	89
soda	75	75	85	108	80	87	87
spaghetti sauce	88	62	78	81	77	68	82
sugar	70	65	81	91	70	81	75
toilet paper	80	66	87	83	83	84	86
tortillas	81	73	56	80	81	79	93
waffles	101	72	82	77	98	76	69
yogurt	84	63	82	86	75	80	75

CNT(dataset_group)

56 108

EDA Inferences

- Yearly, Quarterly, Monthly avg. order count shows an increasing trend
- Especially the Q2’18 and Q1’20 has the highest avg. order count increase
- The weekly trend of avg. order count does not show any pattern
- Weekday trend clearly shows that Weekend and Wednesday has the highest avg. order count across years. This can be utilized for running weekend offers to increase sales.
- Mid week sales campaigns can be done to get more sales
- Across quarters Q3 has the highest avg. orders. Q1 & Q2 has Saturday, Sunday & Wednesday with highest avg. orders. Q3 has Monday, Tuesday & Friday with highest avg. orders.
- 2018: Monday & Wednesday has the highest avg. orders, 2019: Sunday, Monday & Thursday has the highest avg. orders, 2020: Almost same across weekdays

EDA Inferences

- Poultry has been ordered highest no. of times while hand soap been the lowest
- Butter has higher sales in May month.
- Lunch meat has higher sales in the month of July
- Cereals has the higher orders during Sunday and Thursday
- Poultry has the highest orders on Friday and Saturday
- All the items has their higher orders relatively on Sunday

Use of Market Basket Analysis (Association Rules) -->Write Something about the association rules and its relevance in this case

Association rules in market basket analysis are used to identify relationships between items in large datasets which relate to consumer transactions. The market basket analysis is to identify patterns of items that frequently co-occur in transactions

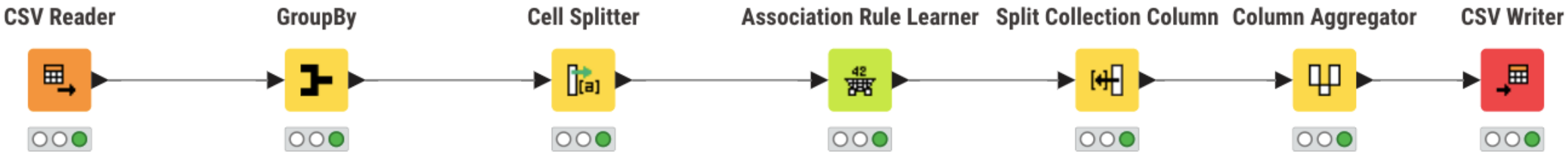
Parameters in Association Rules:

Itemset: These are groups of items that appear together in transactions
Support: How often an item or a group of items appears in your data set.
Confidence: If you have item A, how often item B also appears with it.
Lift: How much more likely item B is bought with item A compared to just buying item B alone.

Relevance:

Association rules in market basket analysis provide valuable insights into customer behaviour, helping businesses to make decisions that enhance customer satisfaction/experience and boost sales. Some of the benefits of Association rule are Cross-Selling and Up-Selling, Store Layout and Shelf Placement, Inventory Management, Targeted Marketing and Promotions

KNIME workflow image



Write about threshold values of Support and Confidence

Support Threshold:

It is a user-defined minimum value that an itemset must meet or exceed to be considered frequent. This threshold helps in filtering out itemset that are not significant or are too rare to be of interest.

In this dataset, the support threshold is set at 0.05 (or 5%), then those itemset that appear in at least 5% of all transactions will be considered.

Confidence Threshold:

It is a user-defined minimum value that an association rule must meet or exceed to be considered strong or reliable.

In this dataset, the confidence threshold is set at 0.5 (or 50%), then only those rules where the itemset setB appears in at least 50% of the transactions containing setA will be considered.

Associations Identified Put the associations in a tabular manner --> Explain about support, confidence, & lift values that are calculated

Rule	Support	Confidence	Lift
{eggs, ice cream, pasta} --> {paper towels}	0.055	0.649	1.791
{paper towels, eggs, ice cream} --> {pasta}	0.055	0.643	1.731
{bagels, cereals, sandwich bags} --> {cheeses}	0.051	0.674	1.726
{yogurt, toilet paper, aluminum foil} --> {juice}	0.050	0.640	1.700
{yogurt, poultry, aluminum foil} --> {mixes}	0.051	0.630	1.678

Association Rule: {eggs, ice cream, pasta} --> {paper towels}
This means that when customers buy eggs, ice cream, and pasta together, they are also likely to buy paper towels.

Support: Measures how frequently the items in the rule appear together in the dataset.
Support: 0.055
Formula: (No. of transactions containing both setA & setB/Total no. of transactions)
This means that 5.5% of all the transactions in the dataset include the items {eggs, ice cream, pasta, paper towels} together.

Confidence: Measures how often the rule {eggs, ice cream, pasta} --> {paper towels} is found to be true.
Confidence: 0.649
Formula: (No. of transactions containing both setA & setB/No. of transactions containing setA)
This means that 64.9% of the transactions that include eggs, ice cream, and pasta also include paper towels.

Lift: Indicates how much more likely the consequent (paper towels) is given the antecedent (eggs, ice cream, pasta) than it would be by chance.
Lift: 1.791
Formula: (Confidence of setA & setB/Support of setB)
This means that buying eggs, ice cream, and pasta together makes it 1.791 times more likely that the customer will also buy paper towels compared to the likelihood of buying paper towels.

Suggestion of Possible Combos with Lucrative Offers --> Write recommendations --> Make discount offers or combos (or buy two get one free) based on the associations and your experience

Combo 1: Eggs, Ice Cream, Pasta, and Paper Towels
Offer: Buy eggs, ice cream, and pasta, get a discount on paper towels.
Discount: Buy any two items from eggs, ice cream, or pasta and get 50% off on paper towels.
Recommendation: The purchase of paper towels with a discount can increase overall sales.

Combo 2: Paper Towels, Eggs, Ice Cream, and Pasta
Offer: Buy paper towels, eggs, and ice cream, get pasta for free.
Discount: Purchase paper towels, eggs, and ice cream together and receive a pack of pasta for free.
Recommendation: Offering pasta for free can make customers to buy the initial three items.

Combo 3: Bagels, Cereals, Sandwich Bags, and Cheeses
Offer: Buy bagels, cereals, and sandwich bags, get cheeses at 25% off.
Discount: Purchase bagels, cereals, and sandwich bags together to receive 25% off on cheeses.
Recommendation: This discount can boost the sale of cheeses.

Combo 4: Yogurt, Toilet Paper, Aluminium Foil, and Juice
Offer: Buy yogurt, toilet paper, and aluminium foil, get juice at 20% off.
Discount: Purchase yogurt, toilet paper, and aluminium foil together and get 20% off on juice.
Recommendation: The discount can drive sales of juice along with the primary items.

Combo 5: Yogurt, Poultry, Aluminium Foil, and Mixes
Offer: Buy yogurt, poultry, and aluminium foil, get a discount on mixes.
Discount: Purchase yogurt, poultry, and aluminium foil together to receive 30% off on baking mixes.
Recommendation: Offering a discount on mixes can promote the sale of all related products.

Thank You