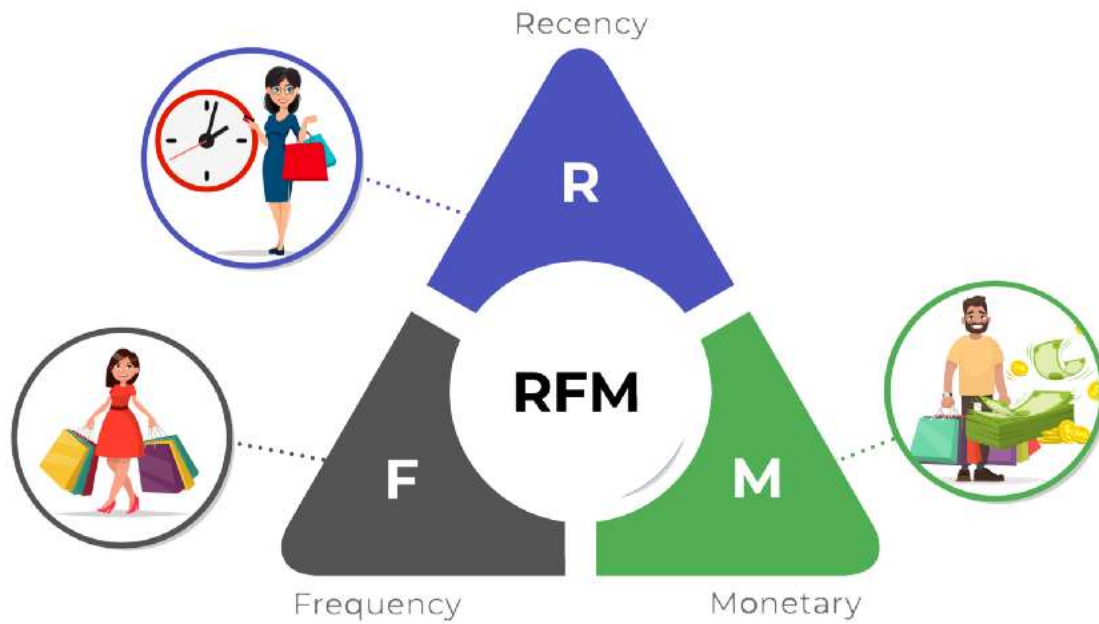


Marketing and Retail Analytics Project



Contents:

RFM analysis:

Agenda and Executive summary

Data summary

Uni-variate, Bi-variate and Multivariate analysis

Yearly analysis, Monthly and Quarterly analysis of the data.

Introduction to RFM

KNIME Workflow

RFM Analysis [BEST CUSTOMERS, LOYAL CUSTOMERS, LOST CUSTOMERS AND CUSTOMERS VERGE OF CHURNING]

Inferences from RFM analysis and identified segments.

Market Basket Analysis:

Agenda and Executive summary

Data summary

Yearly analysis, Monthly and Quarterly analysis of the data.

Use of MBA and Association Rules.

KNIME Workflow

Threshold[VALUES OF SUPPORT AND CONFIDENCE]

Recommendations and Suggestions.

Tableau Link:

https://public.tableau.com/app/profile/pavithra.devi1692/viz/MRAProject_17086943176170/Qquarterlyanalysis

RFM Analysis

Agenda:

To find the key steps in analyzing the company's data to uncover buying patterns, segment customers, and recommend personalized marketing strategies.

Executive Summary: Automobile Parts Manufacturing Company Data Analysis

This dataset, spanning three years and consisting of 2747 transactions, offers a comprehensive view into the operations of an automobile parts manufacturing company. The core columns, including ORDERNUMBER, QUANTITYORDERED, PRICEEACH, ORDERDATE, and SALES, provide crucial information for analyzing sales patterns and customer behavior. The dataset reflects a diverse range of products categorized by PRODUCTLINE, with sales varying across different DEALSIZE segments. Customer-centric details such as CONTACTLASTNAME, CONTACTFIRSTNAME, and DAYS SINCE LASTORDER contribute to understanding purchasing dynamics. The data shows a predominant customer, "Euro Shopping Channel," with distinct buying patterns.

Notable findings include a mean QUANTITYORDERED of 35 items per order and an average sale amount of \$3553. SALES exhibit a wide range, indicating varying order sizes. The dataset also captures the geographical spread of customers, with the majority located in the USA. Exploratory data analysis can uncover further insights into product preferences, seasonal trends, and potential marketing strategies.

Key considerations for the company include leveraging DAYS SINCE LASTORDER to enhance customer retention strategies, tailoring marketing approaches based on identified customer segments, and optimizing inventory based on product popularity. In conclusion, this dataset presents a valuable opportunity for the company to refine its business strategies, foster customer engagement, and maximize profitability.

Data Summary:

The dataset appears well-structured and clean with 2747 rows and 20 columns, containing diverse information about orders, products, and customers. The absence of missing/null values, duplicates, and the broad range in DAYS_SINCE_LASTORDER contribute to the dataset's reliability for further analysis and insights generation.

The dataset comprises various types of columns, including:

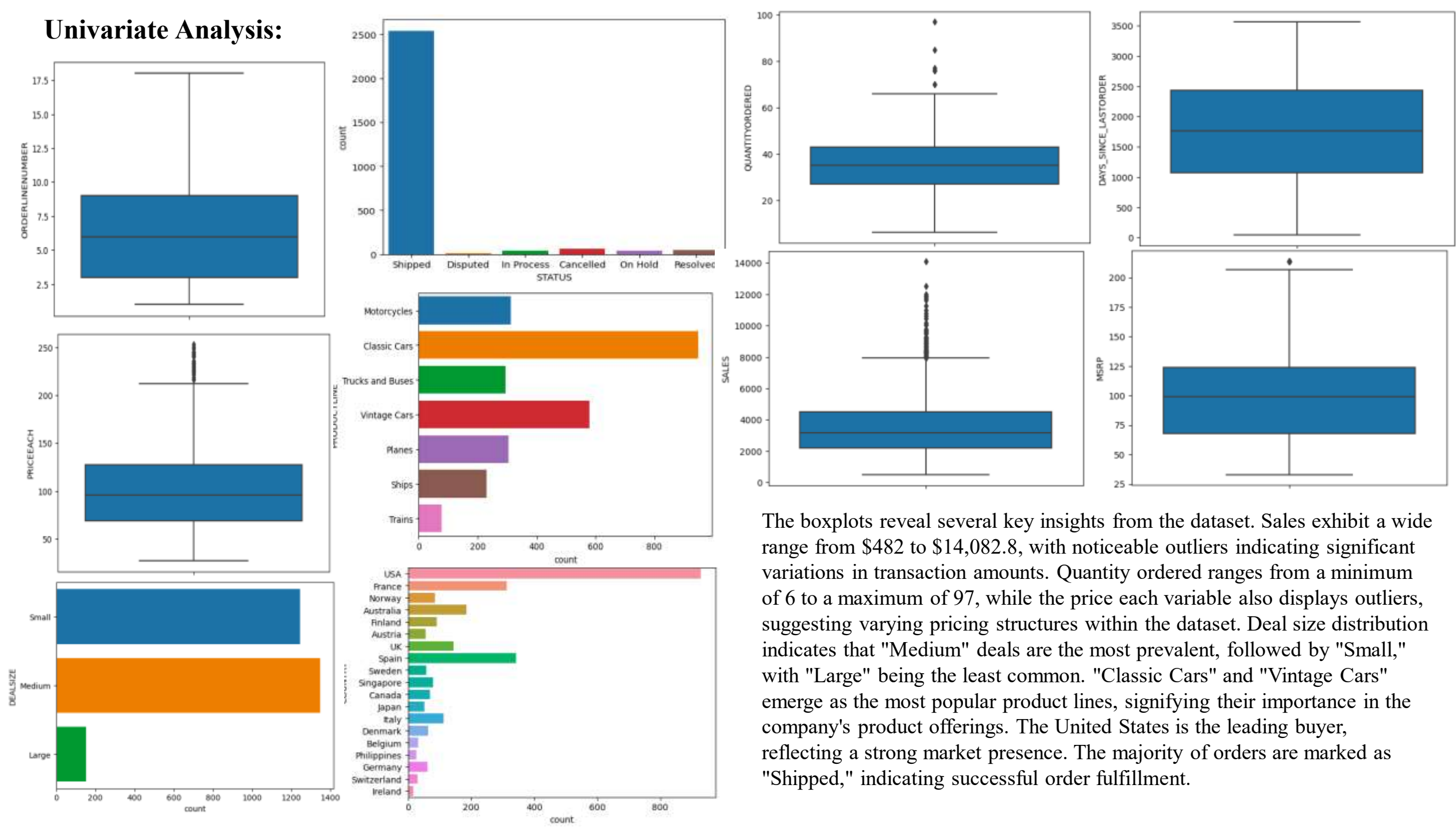
Numeric Columns: ORDERNUMBER, QUANTITYORDERED, PRICEEACH, ORDERLINENUMBER, SALES, ORDERDATE, DAYS_SINCE_LASTORDER, MSRP.

Object/String Columns: STATUS, PRODUCTLINE, PRODUCTCODE, CUSTOMERNAME, PHONE, ADDRESSLINE1, CITY, POSTALCODE, COUNTRY, CONTACTLASTNAME, CONTACTFIRSTNAME, DEALSIZE.

Assuming that details such as CUSTOMERNAME, PHONE, and ADDRESS accurately represent the customers associated with each order, allowing for comprehensive customer analysis.

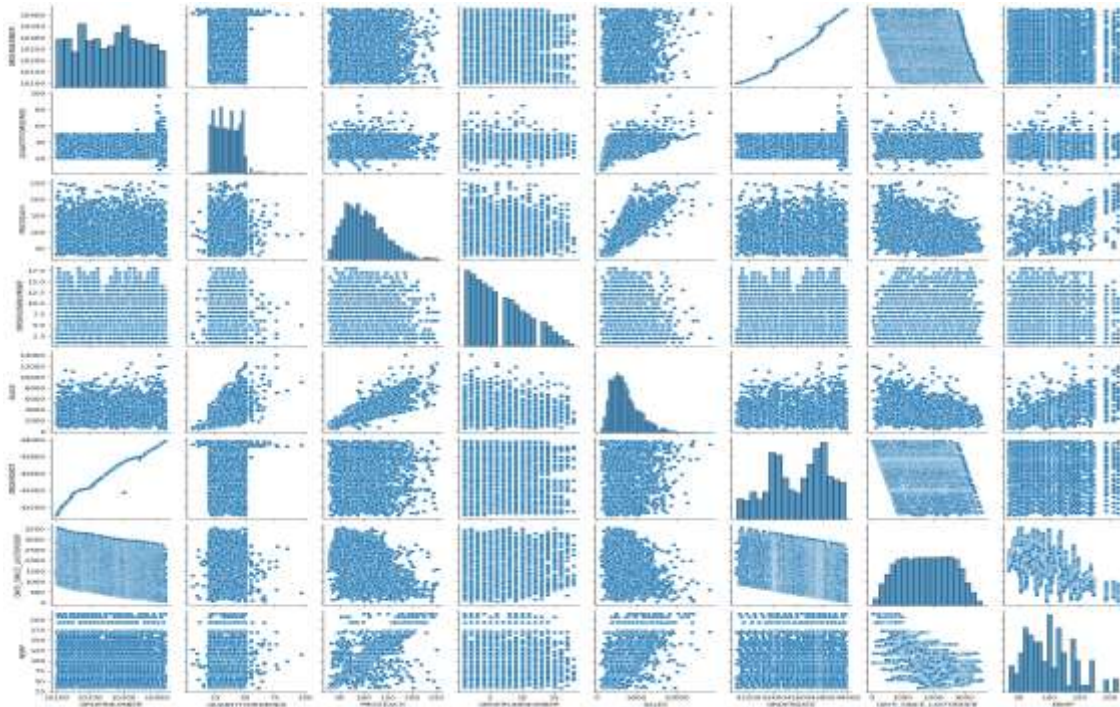
The top frequencies in the dataset shed light on key patterns and preferences. "Shipped" emerges as the dominant order status, implying that the majority of transactions have been successfully fulfilled. "Classic Cars" stands out as the most prevalent product line, suggesting its popularity among customers. The product with the code "S18_3232" and the customer name "Euro Shopping Channel" are frequently occurring, indicating notable product and customer preferences. Additionally, the repeated occurrences of specific values in contact details and location columns suggest a concentration of orders from a dominant customer, potentially located in "Madrid," "USA." "Medium" emerges as the most frequent deal size, indicating a significant proportion of orders fall into this category. These insights can guide the company in tailoring marketing strategies and managing inventory based on the identified popular products and customer segments.

Univariate Analysis:



The boxplots reveal several key insights from the dataset. Sales exhibit a wide range from \$482 to \$14,082.8, with noticeable outliers indicating significant variations in transaction amounts. Quantity ordered ranges from a minimum of 6 to a maximum of 97, while the price each variable also displays outliers, suggesting varying pricing structures within the dataset. Deal size distribution indicates that "Medium" deals are the most prevalent, followed by "Small," with "Large" being the least common. "Classic Cars" and "Vintage Cars" emerge as the most popular product lines, signifying their importance in the company's product offerings. The United States is the leading buyer, reflecting a strong market presence. The majority of orders are marked as "Shipped," indicating successful order fulfillment.

Bi-variate and Multi-variate Analysis:



The heatmap values indicate the correlation coefficients between different pairs of variables in the dataset.

1. Sales and Quantity Ordered (0.55):

- As the quantity ordered increases, sales tend to increase, indicating a logical and expected correlation.

2. Price Each and MSRP (0.78):

- Implies that the actual price each item is sold for tends to align closely with the suggested retail prices.

3. MSRP and Sales (0.63):

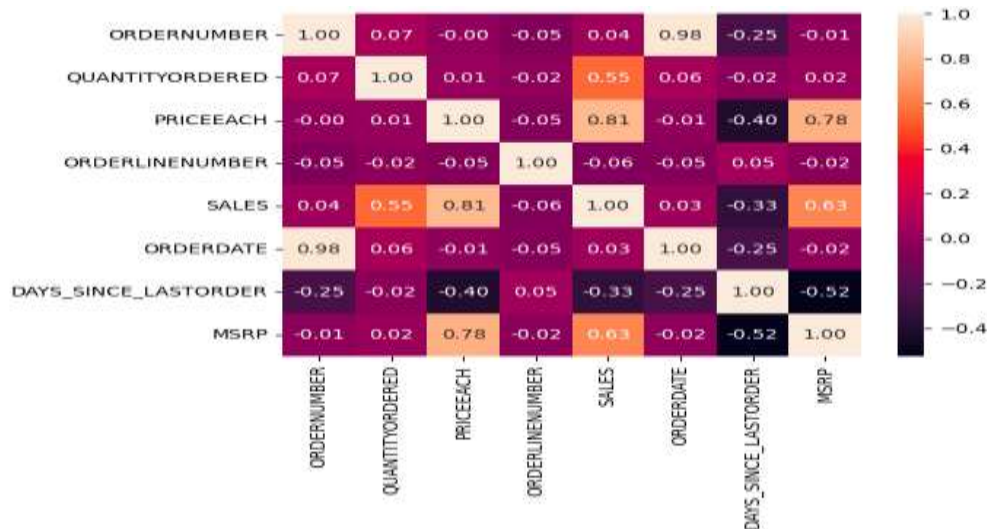
- Higher manufacturer-suggested retail prices may contribute to higher overall sales.

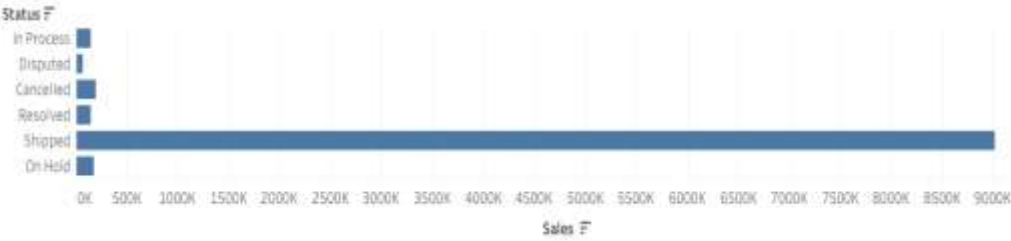
4. Price Each and Sales (0.81):

- As the price of each item increases, total sales also tend to increase substantially.

5. Order Number and Order Date (0.98):

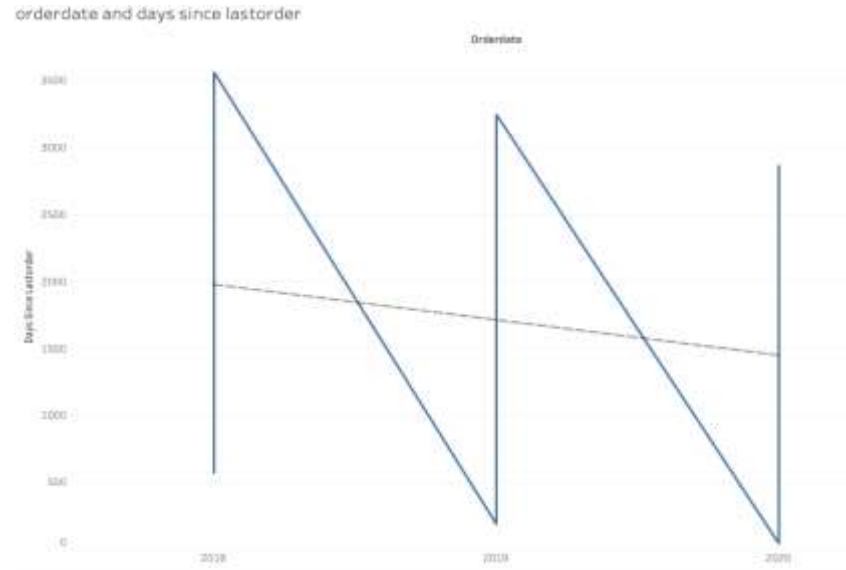
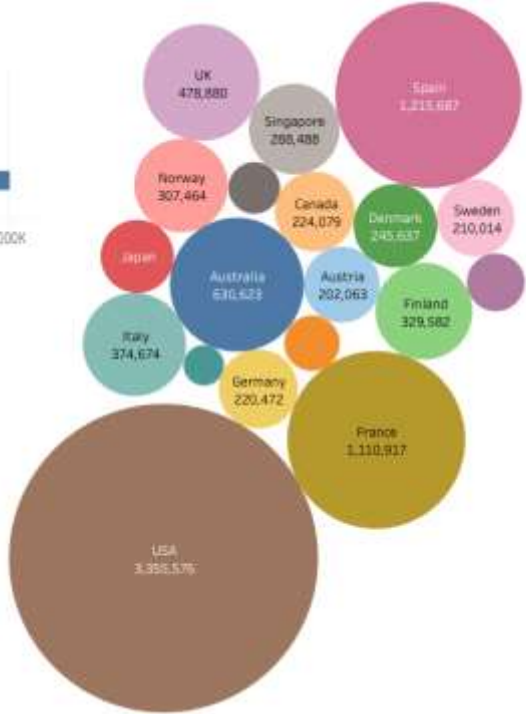
- The very high correlation of 0.98 between Order Number and Order Date implies a nearly perfect positive relationship. This suggests that as the order number increases, the order date also consistently increases, indicating a chronological order of transactions.





Productline

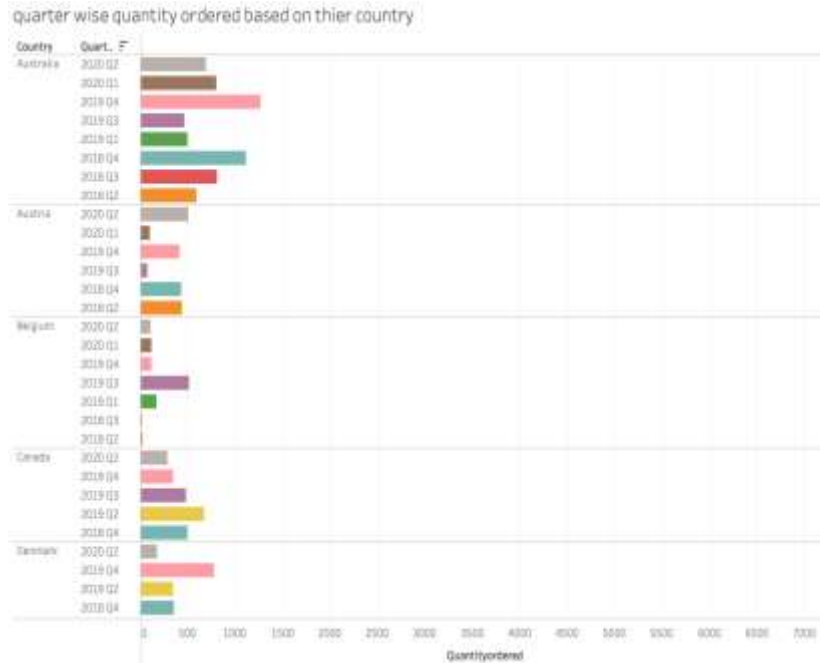
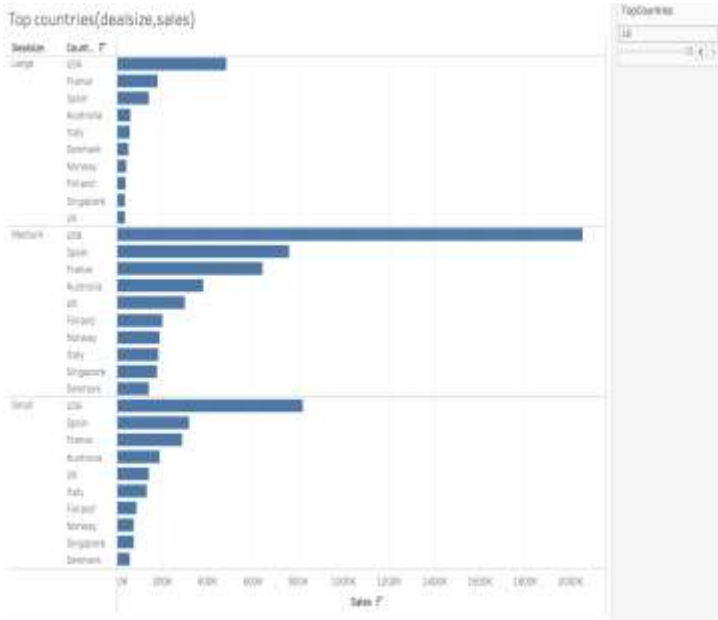
Classic Cars	3,842,869
Vintage Cars	1,806,676
Trucks and Buses	1,111,559
Motorcycles	1,103,512
Planes	969,323
Ships	700,039
Trains	226,243



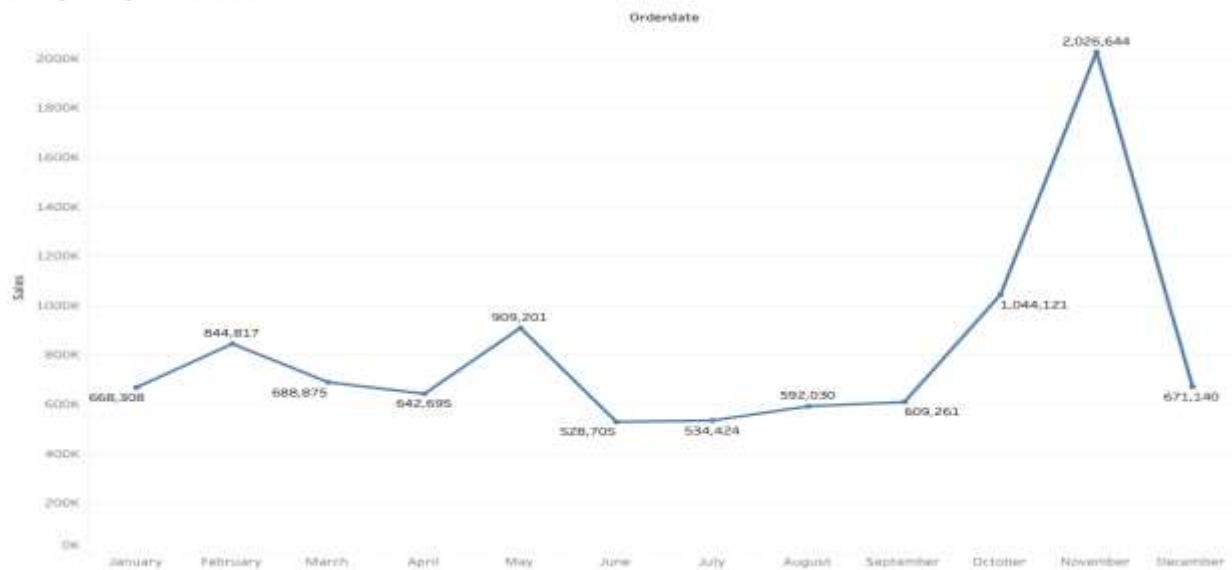
Classic Cars and Vintage Cars collectively account for approximately 57.88% of the total sales.

Examining the sales figures, the United States emerges as the top contributor with sales totaling \$3,355,576, followed by Spain and Italy in succession.

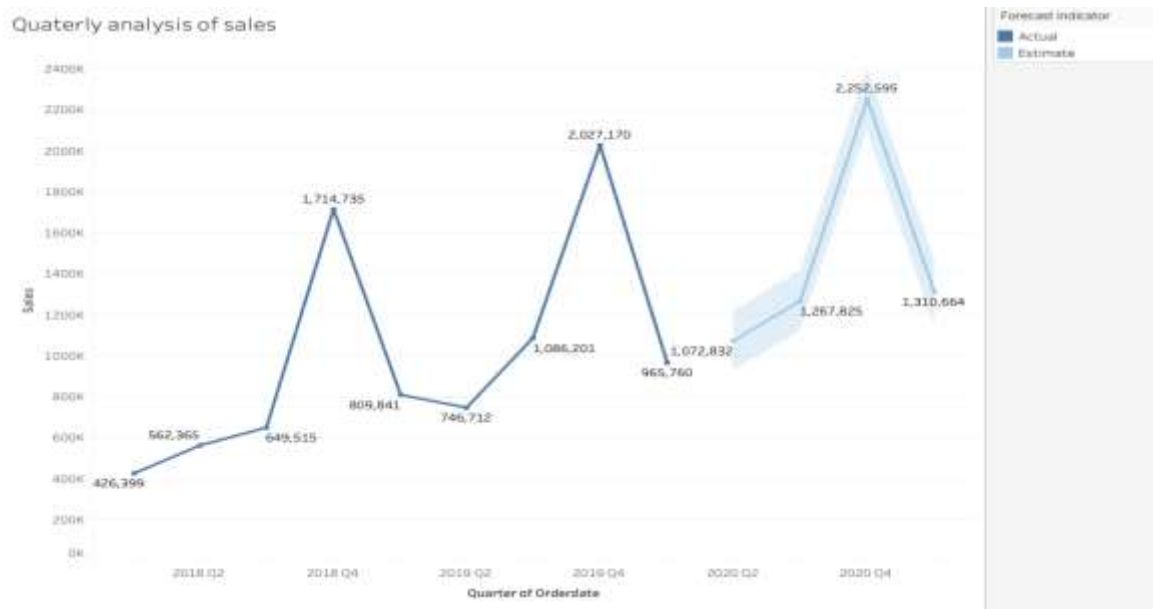
Analyzing the highest sales amounts considering deal size, the leading countries in the first three positions are the United States, France, and Spain.



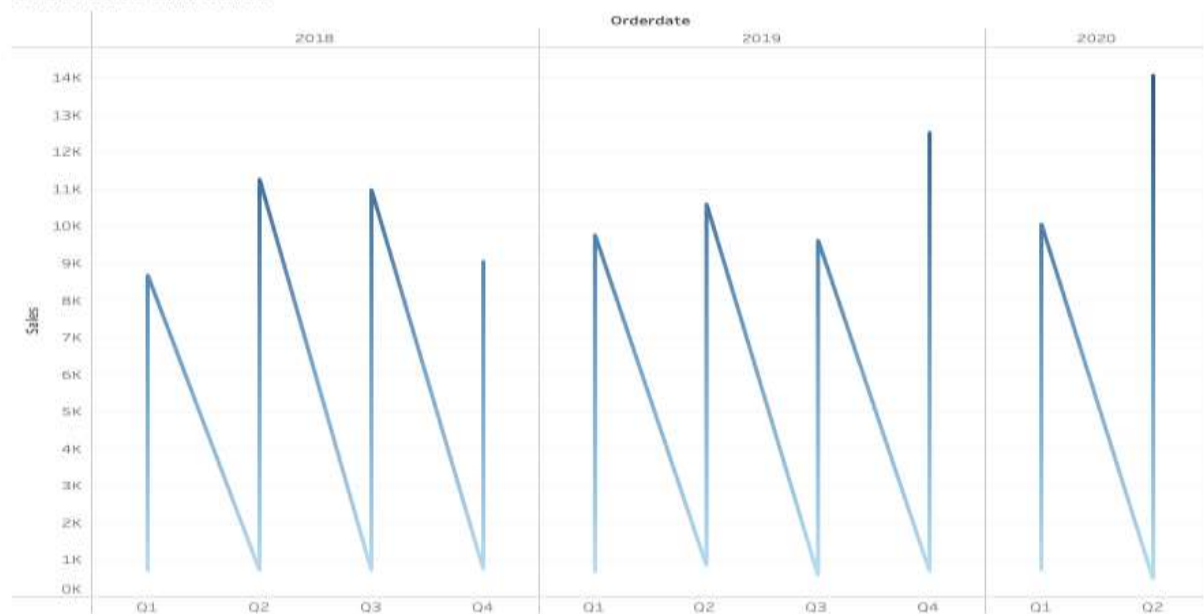
Yearly analysis of sales



Quarterly analysis of sales

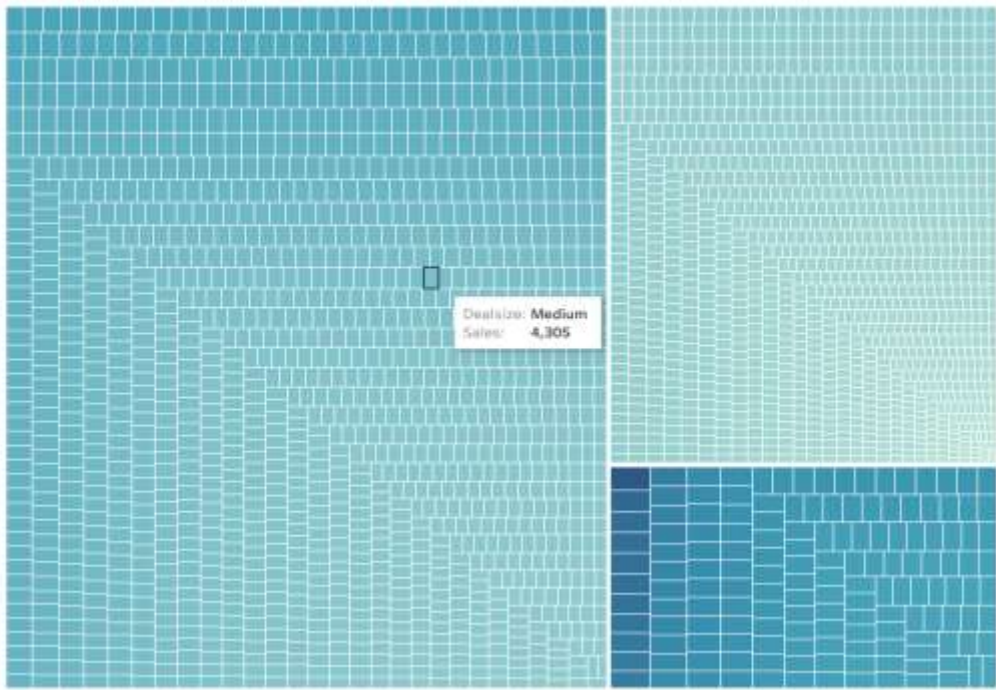


Orderdate and sales



Upon analyzing the graph, it's evident that the highest sales occurred during the month of November, with a noticeable smaller peak observed in February and May. Further examination of the quarterly sales analysis reveals a clear seasonality trend, notably with the quarter 4 consistently exhibiting the highest sales.

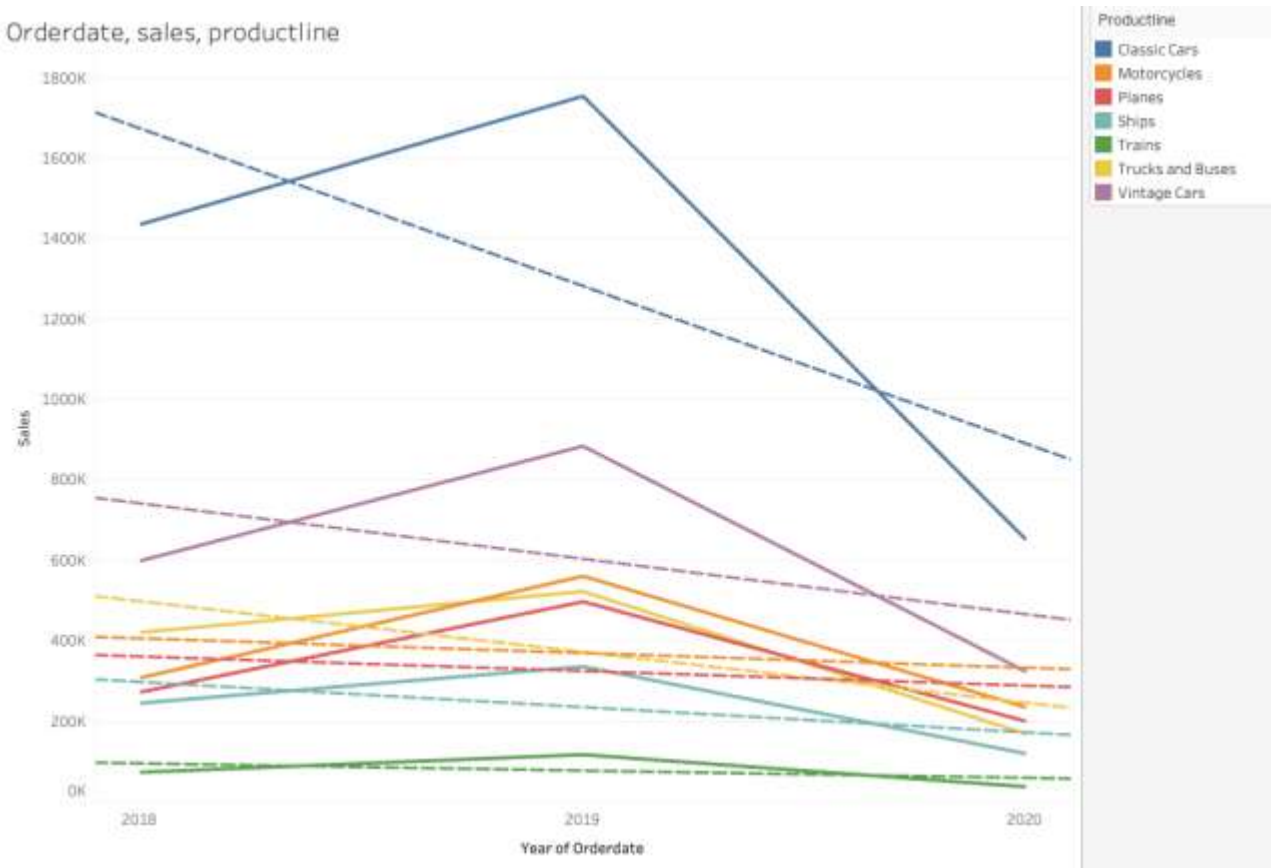
Deal size and sales



Deal size	Productline						
	Classic Cars	Motorcycles	Planes	Ships	Trains	Trucks and Buses	Vintage Cars
Medium	18,746	5,651	5,207	4,262	1,052	6,696	9,595
Small	10,198	4,562	4,952	3,727	1,615	3,525	9,446
Large	4,429	867	477		45	358	1,018

Upon evaluating deal sizes, it is evident that the highest sales are associated with products categorized as having a medium deal size. Specifically, Classic Cars with a medium deal size lead in sales, followed by Motorcycles, also categorized under a medium deal size, and then Planes.

Order date, sales, productline



Upon examining the order date, the sales factor, and the product line from the graph, it is evident that the lowest sales are associated with the product lines of trains, ships, and planes.

Introduction to Customer Segmentation using RFM Analysis:

In the realm of customer relationship management, understanding and categorizing customers based on their purchasing behavior is crucial for personalized marketing strategies. One effective method for segmentation is RFM analysis, a technique that evaluates customer activity using three key metrics: Recency, Frequency, and Monetary Value. Recency refers to the time elapsed since the last customer transaction, Frequency indicates the number of transactions within a specific period, and Monetary Value represents the total amount spent by the customer.

In our RFM analysis, we consider these three parameters to classify customers into distinct segments, allowing for targeted and tailored marketing approaches. The assumptions made include treating a shorter recency period, a higher frequency, and a greater monetary value as indicators of more engaged and valuable customers. To operationalize this analysis, a KNIME workflow is employed, visually showcasing the step-by-step process of RFM segmentation.

Dividing customer segments into four distinct groups:

1.Champions:

1. High Recency (recent purchases)
2. High Frequency (frequent purchases)
3. High Monetary Value (high spending)
4. This segment comprises the most engaged and valuable customers.

2.Potential Loyalists:

1. High Recency (recent purchases)
2. High Frequency (frequent purchases)
3. Low to Moderate Monetary Value
4. Customers in this segment show consistent engagement but may not be the highest spenders.

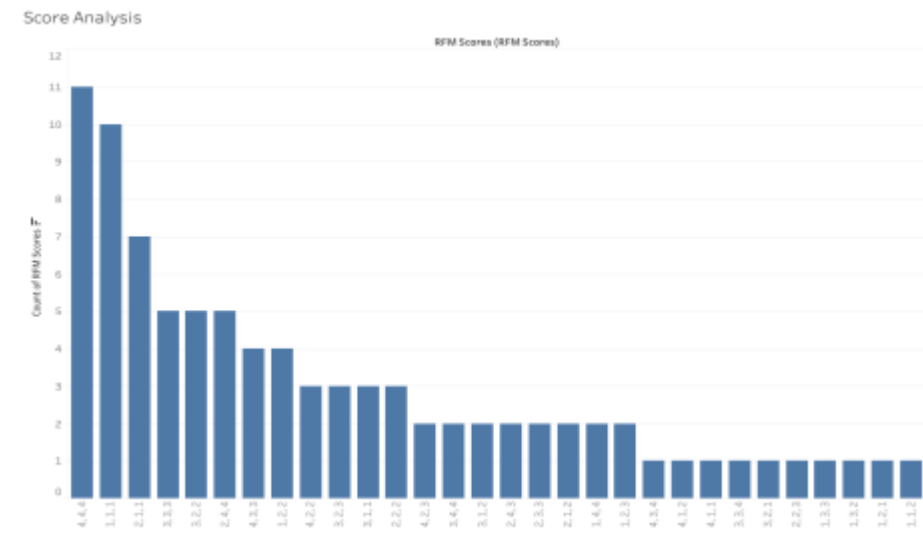
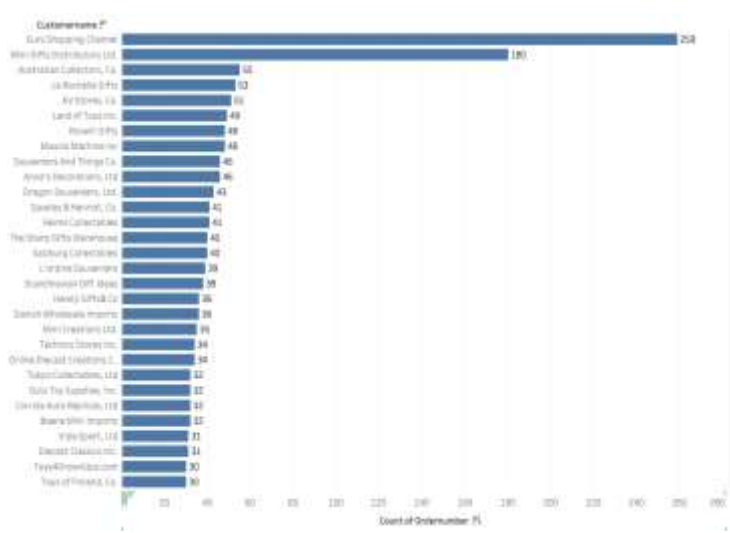
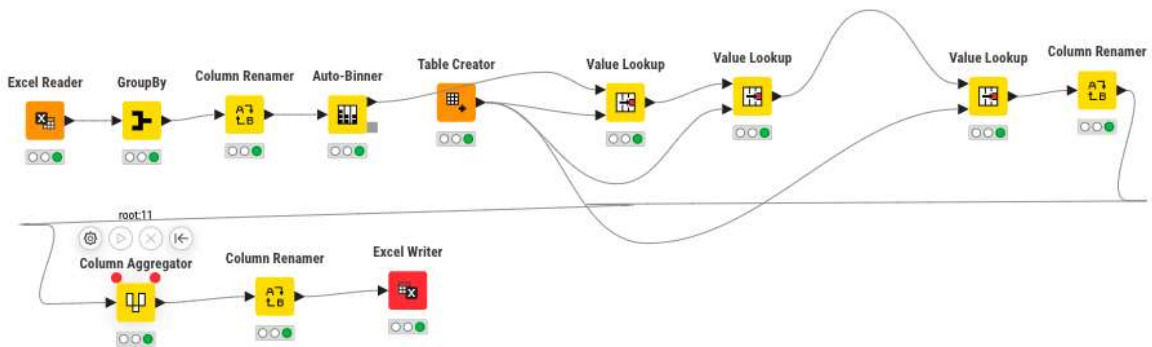
3.At-Risk Customers:

1. Low Recency (not recent purchases)
2. High Frequency (frequent purchases)
3. High Monetary Value
4. These customers were previously engaged but have become less active recently, indicating a risk of churn.

4.Hibernating Customers:

1. Low Recency (not recent purchases)
2. Low Frequency (infrequent purchases)
3. Low to Moderate Monetary Value
4. This segment includes customers who have not engaged recently and may require targeted efforts to re-activate.

In business terms, the ORDERNUMBER serves as a unique identifier for each order, and its count can be interpreted as the frequency of orders. The SALES column, representing the monetary value for each transaction, is considered as the financial aspect of the customer's engagement. Recency, a crucial factor in customer behavior, is monitored through the "days since last order" column. To effectively categorize customers, four distinct bins are created for Recency, Frequency, and Monetary aspects using percentiles (0.0, 0.25, 0.5, 0.75, 1.0). This approach translates to four segments: High(4), Medium(3), Low(2), and Churn(1). These segments aid businesses in strategically classifying and understanding their customer base, facilitating targeted marketing, retention efforts, and tailored approaches based on the recency, frequency, and monetary value of customer transactions.



Recency Sc..	Frequency ..	Monetary Score				Grand T..
		1	2	3	4	
1	1	10	1			11
	2	1	4	2		7
	3		1	1		2
	4				2	2
2	1	7	2			9
	2		3	1		4
	3			2		2
	4			2	5	7
3	1	3	2			5
	2	1	5	3		9
	3			5	1	6
	4				2	2
4	1	1	1			2
	2		3	2		5
	3			4	1	5
	4				11	11
Grand Total		23	22	22	22	89

➡ At-risk customers

 Loyal/Medium Customers

➡ Active/Best Customers

[illegible]

➔ Output of the segmentation.
The table head has the original columns + the recency, frequency and monetary scores and the concatenated RFM scores.

Best Customers

Customername	COUNTRY (..	
Euro Shopping Channel	Spain	912,294
Mini Gifts Distributors Ltd.	USA	654,858
Australian Collectors, Co.	Australia	200,995
La Rochelle Gifts	France	180,125
Land of Toys Inc.	USA	164,069
The Sharp Gifts Warehouse	USA	160,010
Anna's Decorations, Ltd	Australia	153,996
Souvenirs And Things Co.	Australia	151,571
Salzburg Collectables	Austria	149,799
Online Diecast Creations ..	USA	131,685
Technics Stores Inc.	USA	120,783

Loyal Customers

Customername	COUNTRY (..	
Euro Shopping Channel	Spain	912,294
Mini Gifts Distributors Ltd.	USA	654,858
Australian Collectors, Co.	Australia	200,995
La Rochelle Gifts	France	180,125
Land of Toys Inc.	USA	164,069
The Sharp Gifts Warehouse	USA	160,010
Anna's Decorations, Ltd	Australia	153,996
Souvenirs And Things Co.	Australia	151,571
Salzburg Collectables	Austria	149,799
Reims Collectables	France	135,043
Scandinavian Gift Ideas	Sweden	134,259
Online Diecast Creations ..	USA	131,685
Diecast Classics Inc.	USA	122,138
Technics Stores Inc.	USA	120,783
Corrida Auto Replicas, Ltd	Spain	120,615
Tokyo Collectables, Ltd	Japan	120,563
UK Collectables, Ltd.	UK	118,008
Baane Mini Imports	Norway	116,599
Suominen Souvenirs	Finland	113,961
Toys of Finland, Co.	Finland	111,250
Oulu Toy Supplies, Inc.	Finland	104,370
Marta's Replicas Co.	USA	103,080
Gift Depot Inc.	USA	101,895
Heintze Collectables	Denmark	100,596
FunGiftIdeas.com	USA	98,924
Auto Canal Petit	France	93,171
Lyon Souvenirs	France	78,570

Lost Customers

Customername	COUNTRY (..	
Dragon Souvenirs, Ltd.	Singapore	172,990
Rovelli Gifts	Italy	137,956
Toys4GrownUps.com	USA	104,562
Cruz & Sons Co.	Philippines	94,016
Stylish Desk Decors, Co.	UK	88,805
Blauer See Auto, Co.	Germany	85,172
Signal Gift Stores	USA	82,751
Norway Gifts By Mail, Co.	Norway	79,224
Enaco Distributors	Spain	78,412
Marseille Mini Autos	France	74,936
Diecast Collectables	USA	70,860
Alpha Cognac	France	70,488
Australian Collectables, L..	Australia	64,591
Clover Collections, Co.	Ireland	57,756
Gift Ideas Corp.	USA	57,294
Iberia Gift Imports, Corp.	Spain	54,724
Mini Auto Werke	Austria	52,264
Signal Collectibles Ltd.	USA	50,219
Double Decker Gift Stores..	UK	36,019
Bavarian Collectables Imp..	Germany	34,994
Royale Belge	Belgium	33,440
Auto-Moto Classics Inc.	USA	26,479

Customers verge of churning

Customername	COUNTRY (..	
Muscle Machine Inc	USA	197,737
AV Stores, Co.	UK	157,808
Danish Wholesale Imports	Denmark	145,042
Saveley & Henriot, Co.	France	142,874
L'ordine Souvenirs	Italy	142,601
Vida Sport, Ltd	Switzerland	117,714
Handji Gifts& Co	Singapore	115,499
Herkku Gifts	Norway	111,640
Mini Creations Ltd.	USA	108,951
Vitachrome Inc.	USA	88,041
Collectable Mini Designs C..	USA	87,489
Motor Mint Distributors I..	USA	83,682
Mini Caravy	France	80,438
Super Scale Inc.	USA	79,472
Petit Auto	Belgium	74,973
Daedalus Designs Imports	France	69,052
Osaka Souvenirs Co.	Japan	67,605
Online Mini Collectables	USA	57,198
CAF Imports	Spain	49,642
West Coast Collectables C..	USA	46,085
Cambridge Collectables Co.	USA	36,164
Microscale Inc.	USA	33,145

Inferences from RFM Analysis and Identified Segments:

1. High-Value Customers (RFM Score: 4-4-4):

- Recency (R): Customers with a score of 4 demonstrate recent transactions.
- Frequency (F): A score of 4 indicates frequent purchases.
- Monetary (M): Customers scoring 4 exhibit substantial spending.
- Overall RFM Score: "444" segments represent high-value customers who are actively engaged.

2. Potential Loyalists (RFM Score: 3-x-x):

- Recency (R): A moderate score suggests recent engagement.
- Frequency (F): A high score indicates loyalty in past purchases.
- Monetary (M): Moderate to high spending is reflected in the score.
- Overall RFM Score: Segments like "3xx" signify potential loyalists with room for increased spending.

3. At-Risk Customers (RFM Score: 2-4-4):

- Recency (R): Low scores indicate a lapse in engagement.
- Frequency (F): High scores suggest previous loyalty.
- Monetary (M): High spending is reflected in the score.
- Overall RFM Score: "144" segments represent at-risk customers who may benefit from targeted retention efforts.

4. Churned or Inactive Customers (RFM Score: 1-1-1):

- Recency (R): Low scores signify prolonged inactivity.
- Frequency (F): Low scores indicate infrequent purchases.
- Monetary (M): Low spending is reflected in the score.
- Overall RFM Score: "111" segments represent churned or inactive customers who may require re-engagement strategies.

5. Segment-Specific Marketing Strategies:

- High-Value Segment: Develop exclusive offers or loyalty programs for customers in the "444" segment.
- Potential Loyalists: Implement targeted campaigns to encourage increased spending in segments like "3**."
- At-Risk Segment: Deploy retention strategies such as special promotions for customers in the "2**" segment.
- Churned/Inactive Segment: Design reactivation campaigns for customers in the "111" segment.

6. Strategic Decision-Making:

- Resource Allocation: Efficiently allocate marketing resources by focusing on segments with higher potential returns, represented by higher RFM scores.
- Personalized Communication: Tailor communication strategies based on the specific needs and behaviors indicated by RFM scores.
- Retention Planning: Develop targeted retention plans to mitigate the risk of customer churn, particularly in segments with lower RFM scores.

Market Basket Analysis

Agenda:

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.

Executive Summary:

	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
...
20636	25-02-2020	1138	soda
20637	25-02-2020	1138	paper towels
20638	26-02-2020	1139	soda
20639	26-02-2020	1139	laundry detergent
20640	26-02-2020	1139	shampoo

20641 rows x 3 columns

Description of the dataset

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Date	20641	603	08-02-2019	183	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Order_id	20641.0	NaN	NaN	NaN	575.986289	328.557078	1.0	292.0	581.0	862.0	1139.0
Product	20641	37	poultry	640	NaN	NaN	NaN	NaN	NaN	NaN	NaN

	Date	Order_id	Product
10	01-01-2018	1	all- purpose
13	01-01-2018	1	all- purpose
18	01-01-2018	1	dinner rolls
29	01-01-2018	2	waffles
31	01-01-2018	2	hand soap
...
20616	24-02-2020	1137	paper towels
20632	25-02-2020	1138	sandwich bags
20633	25-02-2020	1138	toilet paper
20635	25-02-2020	1138	soda
20636	25-02-2020	1138	soda

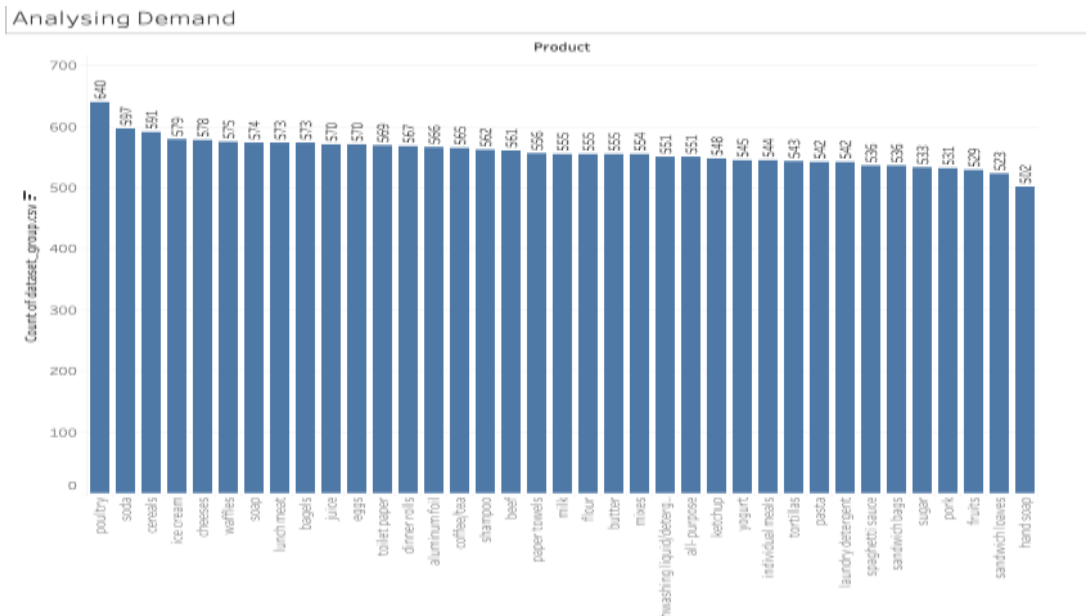
4730 rows x 3 columns

```
<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)
```

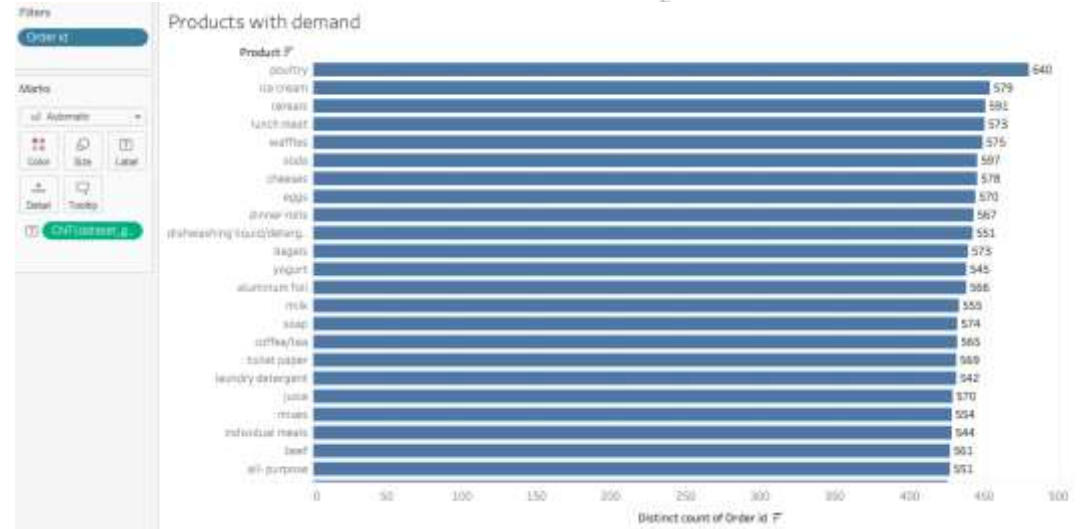
Non null

Duplicate items

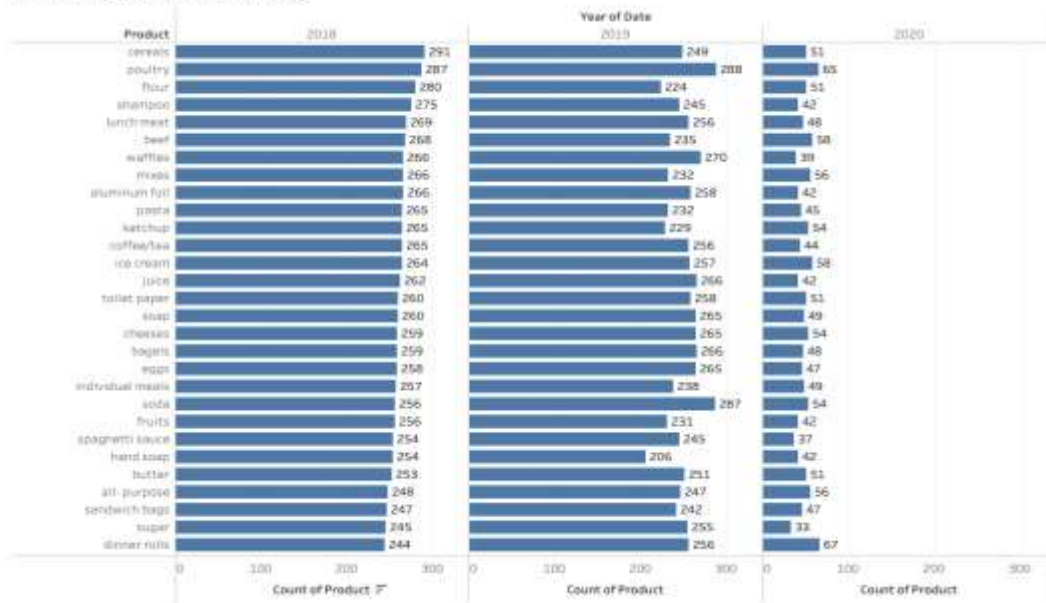
The dataset comprises 20,641 rows and three columns, denoted as date, order ID, and product. Among these columns, there are 37 distinct products, with poultry being the most frequently occurring. A total of 4,730 duplicates have been identified, signifying repeated purchases of products. However, the decision has been made to proceed with the analysis without addressing the duplicates. Examining the data types, the dataset contains two object data types and one integer data type.



Upon analyzing the overall product demand, it is evident that there is significant popularity for items such as poultry, soda, cereals, ice creams, and waffles, indicating a high general demand. However, when the demand is scrutinized based on distinct order IDs, a distinct preference emerges, with a notable demand for poultry, ice cream, cereal, lunch meat, and waffles. Conversely, there is comparatively low demand for products like sugar, fruits, sandwich loaves, hand soaps. Interestingly, when assessed on a per-order basis, it becomes apparent that the demand for products like sugar, pork, flour, sandwich loaves, and hand soap is consistently low.



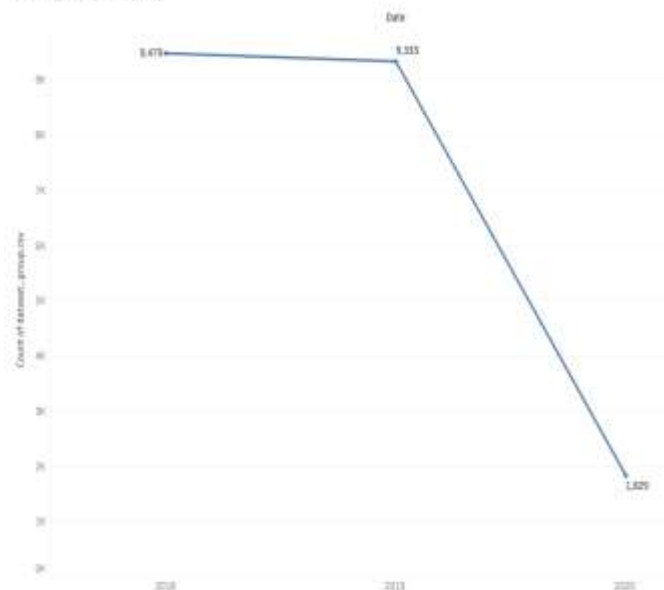
Yearly analysis of the products



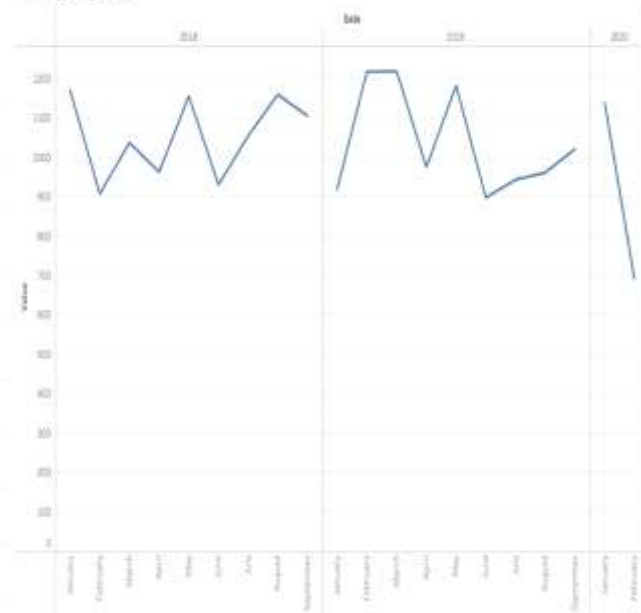
The analysis of product demand over the years reveals distinct patterns and shifts in consumer preferences. In 2018, there was heightened demand for items like cereal, poultry, flour, shampoo, and lunch meat. The subsequent year, 2019, witnessed a shift in demand towards poultry, soda, dishwashing detergent, waffles, and milk. By 2020, consumer preferences evolved further, with the highest demand observed for dinner rolls, poultry, poke, ice cream, and beef. These trends suggest a dynamic market landscape where product popularity can fluctuate, emphasizing the importance of adapting strategies to evolving consumer preferences for effective inventory management and marketing initiatives.

But the demand for poultry is always in the top 5.

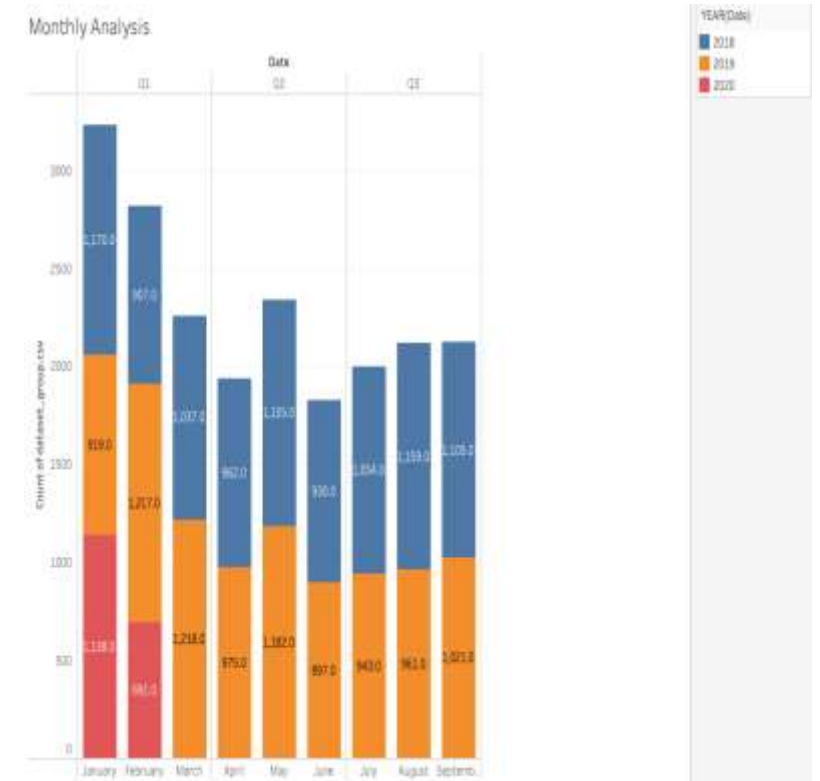
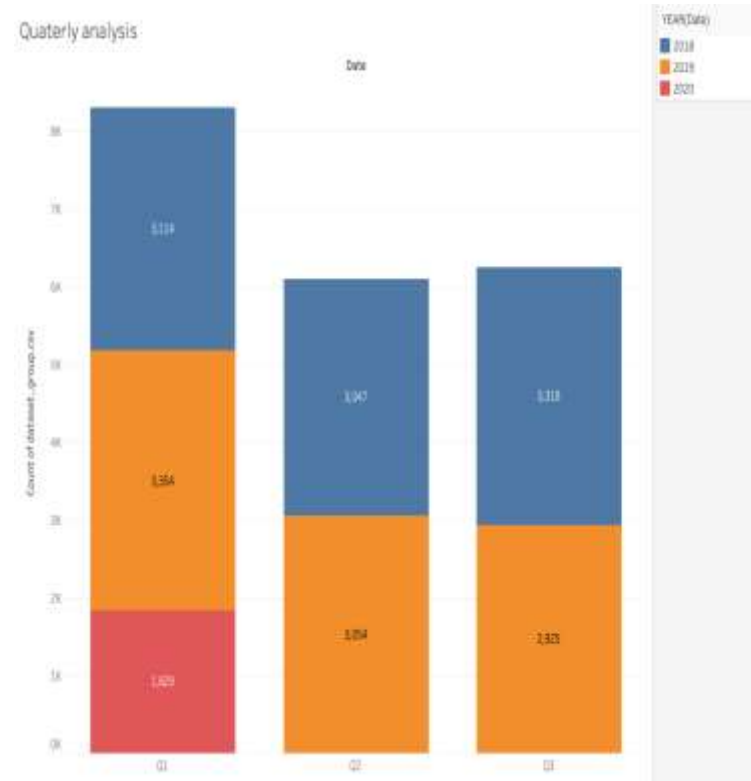
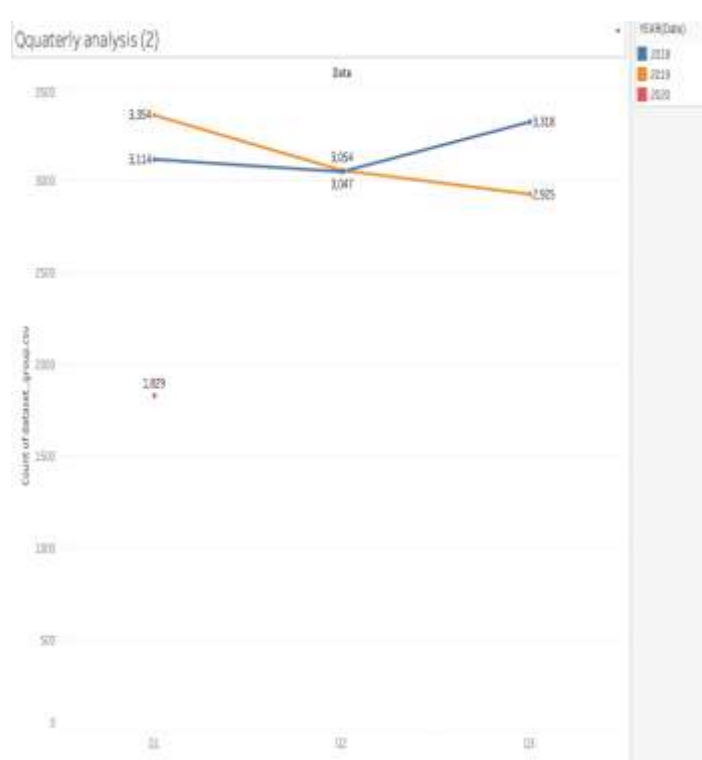
MRA-2(yearly analysis)



Monthly analysis [2]



In the consolidated data, the highest number of orders is observed in 2018, while 2020 registers the lowest count, attributed to the availability of only two months' worth of data. It is likely that orders in 2020 fall within a similar range. Examining the overall order trend, there is a decrease over the years, with 2018 recording the highest number of orders. Furthermore, when assessing the monthly order trends within the given dataset, there is an observed seasonal spike in the number of orders during the month of May. January and February emerge as the months with the highest order activity.



Upon analyzing the orders quarterly, a discernible pattern emerges. Quarter one consistently exhibits the highest number of orders, followed by a decline in quarter two. Subsequently, there is an uptick in orders during quarter three. Specifically, in 2019, the highest number of orders is observed in both quarter one and quarter two. However, during quarter three, 2018 surpasses 2019 in terms of order count, showcasing a distinctive pattern in quarterly order distribution.

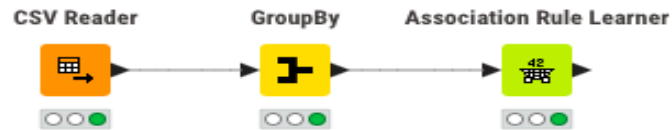
Use of Market Basket Analysis (Association Rules):

Association rules, commonly employed in Market Basket Analysis, play a crucial role in extracting meaningful insights from transactional data. In the context of this grocery store dataset, association rules help uncover relationships between products that are frequently purchased together. This analysis aids in identifying patterns, facilitating the creation of strategic combo offers and targeted promotions.

Relevance in this Case:

Understanding product associations enables the grocery store to optimize product placement, enhance cross-selling, and design effective marketing campaigns. By leveraging association rules, the store can tailor promotional strategies to encourage customers to buy complementary items, ultimately boosting sales and customer satisfaction.

KNIME Workflow Image:



The KNIME workflow showcases the step-by-step process of implementing Market Basket Analysis. This includes data preprocessing, association rule generation, and result visualization, providing a comprehensive view of the analytical workflow.

Threshold Values of Support and Confidence:



The screenshot shows the 'Itemset Mining' dialog box in KNIME. It has four tabs: 'Options', 'Flow Variables', 'Job Manager Selection', and 'Memory Policy'. The 'Options' tab is active. Under 'Itemset Mining', 'Column containing transactions' is set to 'Product', 'Minimum support (0-1)' is 0.02, and 'Underlying data structure' is 'ARRAY'. Under 'Output', 'Itemset type' is 'CLOSED' and 'Maximal itemset length' is 10. Under 'Association Rules', 'Output association rules' is checked and 'Minimum confidence' is 0.8. At the bottom are 'OK', 'Apply', and 'Cancel' buttons.

In this KNIME workflow, the chosen parameter values for support and confidence are crucial in determining the relevance and significance of the association rules.

- Support (0.02):

- The support threshold of 0.02 implies that only associations occurring in at least 2% of the transactions are considered. This ensures that the rules generated are based on reasonably frequent patterns.

- Confidence (0.8):

- The confidence threshold of 0.8 indicates that only rules with a confidence level of 80% or higher are considered. This ensures that the identified associations are highly reliable and likely to occur, providing actionable insights for the grocery store's marketing strategies.

These threshold values strike a balance between capturing meaningful associations and avoiding an overwhelming number of rules. The chosen values reflect a focus on strong associations that can drive targeted promotional efforts.

Conclusion:

Market Basket Analysis with association rules, as implemented in this KNIME workflow, empowers the grocery store to make informed decisions regarding product bundling, promotional activities, and overall merchandising strategies, thereby maximizing revenue and enhancing the overall customer shopping experience.

Associations Identified:

<input type="checkbox"/>	#	RowID	Support <i>Number (double)</i>	Confidence <i>Number (double)</i>	Lift <i>Number (double)</i>	Consequent <i>String</i>	implies <i>String</i>	Items <i>Set</i>
<input type="checkbox"/>	1	rule0	0.02	0.852	2.267	mixes	<---	[yogurt,dishwashing liquid/deter...
<input type="checkbox"/>	2	rule1	0.02	0.821	2.136	yogurt	<---	[cheeses,all- purpose,tortillas,...]
<input type="checkbox"/>	3	rule2	0.02	0.821	2.191	beef	<---	[shampoo,fruits,lunch meat,...]
<input type="checkbox"/>	4	rule3	0.02	0.852	2.18	soda	<---	[bagels,pasta,individual meals,...]

1. Support:

- Support measures the proportion of transactions that contain the specified set of items
- Rule 1 (Mixes): Appears in 2% of all transactions.
- **Rule 2 (Yogurt): Occurs in 2% of all transactions.
- Rule 3 (Beef):Present in 2% of all transactions.
- Rule 4 (Soda) Found in 2% of all transactions.
- Rule 5 (Milk): Occurs in 2% of all transactions.

2. Confidence:

- Confidence quantifies how often the rule is true.
- Rule 1 (Mixes): 85.2% of transactions containing antecedent items also have the consequent item.
- Rule 2 (Yogurt): 82.1% confidence indicates that 82.1% of transactions with antecedent items also include yogurt.
- Rule 3 (Beef):82.1% confidence suggests that 82.1% of transactions with antecedent items also include beef.
- Rule 4 (Soda):85.2% confidence signifies that 85.2% of transactions with antecedent items also include soda.
- Rule 5 (Milk):82.1% confidence implies that 82.1% of transactions with antecedent items also include milk.

3. Lift:

- Lift measures how much more likely items are purchased together compared to when they are purchased individually.
- Rule 1 (Mixes): Lift of 2.267 suggests a 126.7% increase in the likelihood of purchasing items in the rule together compared to random chance.
- Rule 2 (Yogurt): Lift of 2.136 indicates a 113.6% increase in the likelihood of purchasing items in the rule together compared to random chance.
- Rule 3 (Beef): Lift of 2.191 implies a 119.1% increase in the likelihood of purchasing items in the rule together compared to random chance.
- Rule 4 (Soda): Lift of 2.18 suggests a 118% increase in the likelihood of purchasing items in the rule together compared to random chance.
- Rule 5 (Milk): Lift of 2.161 indicates a 116.1% increase in the likelihood of purchasing items in the rule together compared to random chance.

These values provide insights into the strength and significance of the identified associations, guiding the understanding of customer behavior and influencing potential marketing strategies.

<input type="checkbox"/>	#	RowID	Support Number (double)	Confidence Number (double)	Lift ↓ Number (double)	Consequent String	Implies String	Items Set
<input type="checkbox"/>	8	rule7	0.02	0.852	2.349	paper towels	←	[eggs,dinner rolls,ice cream,...]
<input type="checkbox"/>	1	rule0	0.02	0.852	2.267	mixes	←	[yogurt,dishwashing liquid/deter...
<input type="checkbox"/>	9	rule8	0.02	0.821	2.265	paper towels	←	[eggs,dinner rolls,poultry,...]
<input type="checkbox"/>	24	rule23	0.023	0.839	2.258	ketchup	←	[tortillas,coffee/tea,juice,...]
<input type="checkbox"/>	18	rule17	0.022	0.833	2.244	pasta	←	[paper towels,dishwashing liquid...
<input type="checkbox"/>	7	rule6	0.02	0.885	2.219	ice cream	←	[paper towels,eggs,dinner rolls,...]
<input type="checkbox"/>	12	rule11	0.021	0.828	2.218	spaghetti sauce	←	[waffles,laundry detergent,mixes,...]
<input type="checkbox"/>	11	rule10	0.021	0.828	2.208	beef	←	[poultry,fruits,hand soap,...]

Suggested Combos with Lucrative Offers:

1. Combo: Paper Towels

- ****Offer:**** Buy a pack of Paper Towels with a special discount.
- ****Recommendation:**** Customers who purchase Eggs, Dinner Rolls, Ice Cream, Pasta, and Lunch Meat have a high likelihood of also buying Paper Towels.

Attract them with a discounted bundle.

2. Combo: Mixes

- ****Offer:**** Special discount on Mixes.
- ****Recommendation:**** Customers who buy Yogurt, Dishwashing Liquid/Detergent, All-Purpose, and Hand Soap are likely to purchase Mixes as well. Create an attractive offer for this combo.

3. Combo: Paper Towels

- ****Offer:**** Bundle deal on Paper Towels.
- ****Recommendation:**** Customers who purchase Eggs, Dinner Rolls, Poultry, Ice Cream, and Pasta are frequently interested in Paper Towels. Offer them a discounted package.

4. Combo: Ketchup

- ****Offer:**** Special promotion on Ketchup.
- ****Recommendation:**** Customers buying Tortillas, Coffee/Tea, Juice, and Soap are often interested in Ketchup. Create a combo offer to boost sales.

5. Combo: Pasta

- ****Offer:**** Buy-two-get-one-free on Pasta.
- ****Recommendation:**** Customers buying Paper Towels, Dishwashing Liquid/Detergent, Eggs, Dinner Rolls, and Ice Cream are likely to purchase Pasta.

Encourage bulk buying with a special offer.

Recommendations and Suggestions:

1. Bundle Offers:

- Implement bundle offers for frequently associated items identified through market basket analysis. For example, offer special discounts when customers purchase items like Paper Towels along with Eggs, Dinner Rolls, Ice Cream, Pasta, or Lunch Meat.

2. Loyalty Programs:

- Introduce loyalty programs to reward customers who consistently buy from the identified associations. Encourage repeat purchases by offering points or discounts for every qualified transaction.

3. Targeted Promotions:

- Create targeted promotions for specific product combinations. For instance, promote Mixes with discounts for customers who frequently buy Yogurt, Dishwashing Liquid/Detergent, All-Purpose, and Hand Soap.

4. Buy-One-Get-One (BOGO) Offers:

- Implement buy-one-get-one-free or buy-two-get-one-free offers for products that show strong associations. This can incentivize customers to purchase more and increase their basket size.

5. Cross-Marketing:

- Leverage cross-marketing opportunities by strategically placing associated items near each other in the store. Encourage impulse purchases by showcasing items that often go together.

6. Customer Engagement:

- Engage customers through targeted communication about special offers and discounts on associated products. Utilize email campaigns, in-store signage, and social media to promote these deals.

7. Personalized Recommendations:

- Implement a personalized recommendation system on the online platform based on the identified associations. This can enhance the customer shopping experience and drive additional sales.

8. Feedback and Adjustment:

- Regularly monitor the performance of implemented strategies and gather customer feedback. Adjust promotions and offers based on the effectiveness and customer response.

9. Data-Driven Decision-Making:

- Continue utilizing data-driven decision-making processes. Regularly analyze market basket data to identify evolving trends and adjust marketing strategies accordingly.

10. Seasonal Promotions:

- Tailor promotions to align with seasonal trends. Create special offers or discounts during holidays or events, taking into consideration the identified associations.

Implementing these recommendations can not only increase revenue through strategic marketing but also enhance customer satisfaction by providing value through targeted promotions and offers. Regularly reviewing and adapting strategies based on market basket analysis insights will contribute to long-term success in maximizing sales opportunities.