MARKETING AND RETAIL ANALYSIS

AUTOMOBILE PARTS MANUFACTURING COMPANY

ANALYSIS DONE BY

- ANAY TIWARI

CONTENT

- 1. Agenda and Executive Summary of the Data

- 2. Exploratory Analysis and Inferences

- 3. Customer Segmentation using RFM Analysis

- 4. Inferences from RFM Analysis and Identified Segments

- 5. Recommendations

1. AGENDA AND EXECUTIVE SUMMARY OF THE DATA

- Problem Statement

- About Data (Info, Shape, Summary, Statistical Summary)

- Assumptions on Data

PROBLEM STATEMENT

- An automobile parts manufacturing company has collected data on transactions for 3 years. They do not have any in-house data science team, thus they have hired you as their consultant. Your job is to use your data science skills to find the underlying buying patterns of the customers, provide the company with suitable insights about their customers, and recommend customized marketing strategies for different segments of customers.

SUMMARY OF THE DATA

- The dataset consists of 2,747 transactions made by customers of an automobile parts manufacturing company.

- The dataset consists of 20 columns, including 2 float, 5 integer, and 13 object data types, providing a diverse range of information about the transactions and customers.

- The dataset is free from duplicate entries. There are no duplicates present in the dataset.

- Customers on average ordered 35 items per transaction, with a minimum of 6 and a maximum of 97.

- The average number of days since the last order is 1,757.09, with a minimum of 42 days and a maximum of 3,562 days.

- The average Manufacturer's Suggested Retail Price (MSRP) of the products is $100.69, with a standard deviation of $40.11.

STATISTICAL SUMMARY

|	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	DAYS_SINCE_LASTORDER	MSRP
count	2747.000000	2747.000000	2747.000000	2747.000000	2747.000000	2747.000000	2747.000000
mean	10259.761558	35.103021	101.098951	6.491081	3553.047583	1757.085912	100.691664
std	91.877521	9.762135	42.042548	4.230544	1838.953901	819.280576	40.114802
min	10100.000000	6.000000	26.880000	1.000000	482.130000	42.000000	33.000000
25%	10181.000000	27.000000	68.745000	3.000000	2204.350000	1077.000000	68.000000
50%	10264.000000	35.000000	95.550000	6.000000	3184.800000	1761.000000	99.000000
75%	10334.500000	43.000000	127.100000	9.000000	4503.095000	2436.500000	124.000000
max	10425.000000	97.000000	252.870000	18.000000	14082.800000	3562.000000	214.000000

ASSUMPTIONS ON THE DATA

- The dataset comprises information on customer orders, each uniquely identified by an ORDERNUMBER.

- It includes transactional details such as the quantity of items purchased, the time since the last order, and the Manufacturer's Suggested Retail Price (MSRP) of the products.

- Duplicate entries are absent, ensuring that each transaction record is distinct and not repeated for the same customer or transaction.

2. EXPLORATORY ANALYSIS AND INFERENCES

PRODUCT LINE ANALYSIS
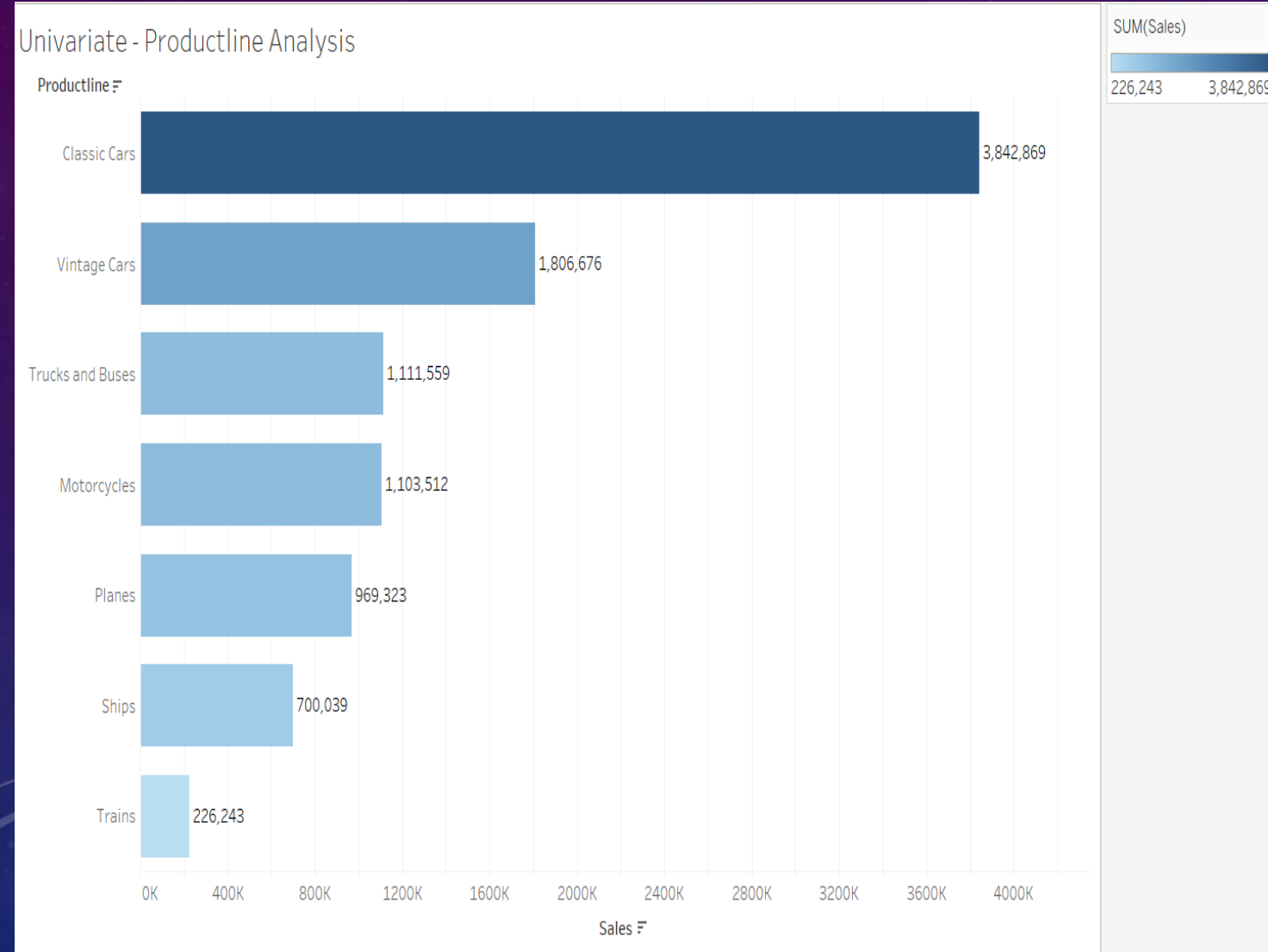
Inference :

1. Classic cars and Vintage cars emerge as top performing product lines, indicating robust demand for these specific vehicle models among customers.
2. Ships, Trucks, Buses and Trains are exhibit comparatively lower counts, hinting at a niche market or specialized clientele for these products categories.

DEAL SIZE ANALYSIS

Inference :

1. The dataset reveals that the bulk of deals fall within the small and medium categories, with counts of 1246 and 1349.
2. The trend suggests that the company predominantly conducts transactions with moderate to smaller deal sizes, emphasizing a focus on transactions of this scale.
3. They should focus on large deal size as well.

Dealsize Analysis
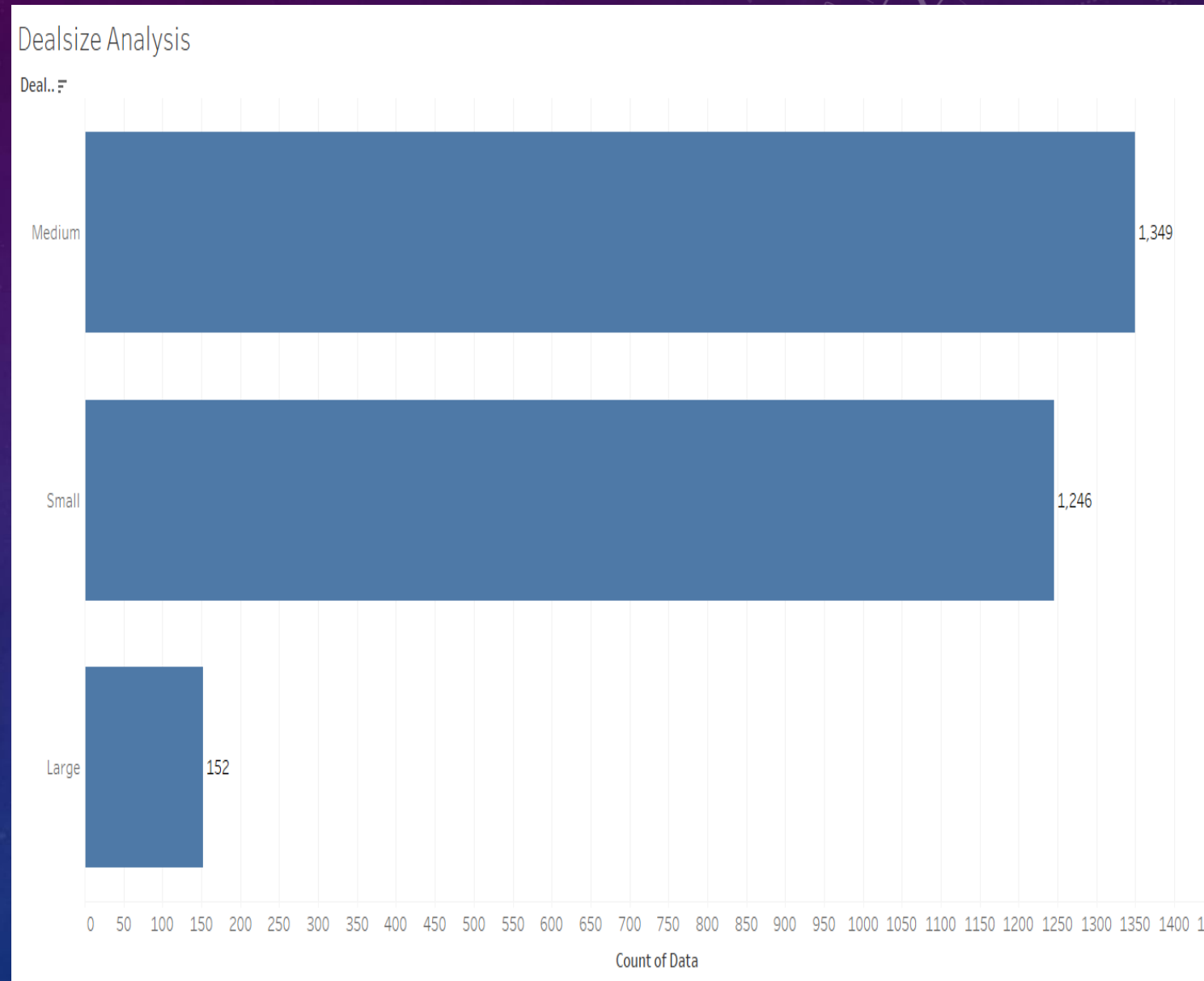
ORDER STATUS ANALYSIS

Univariate Order Status Analysis
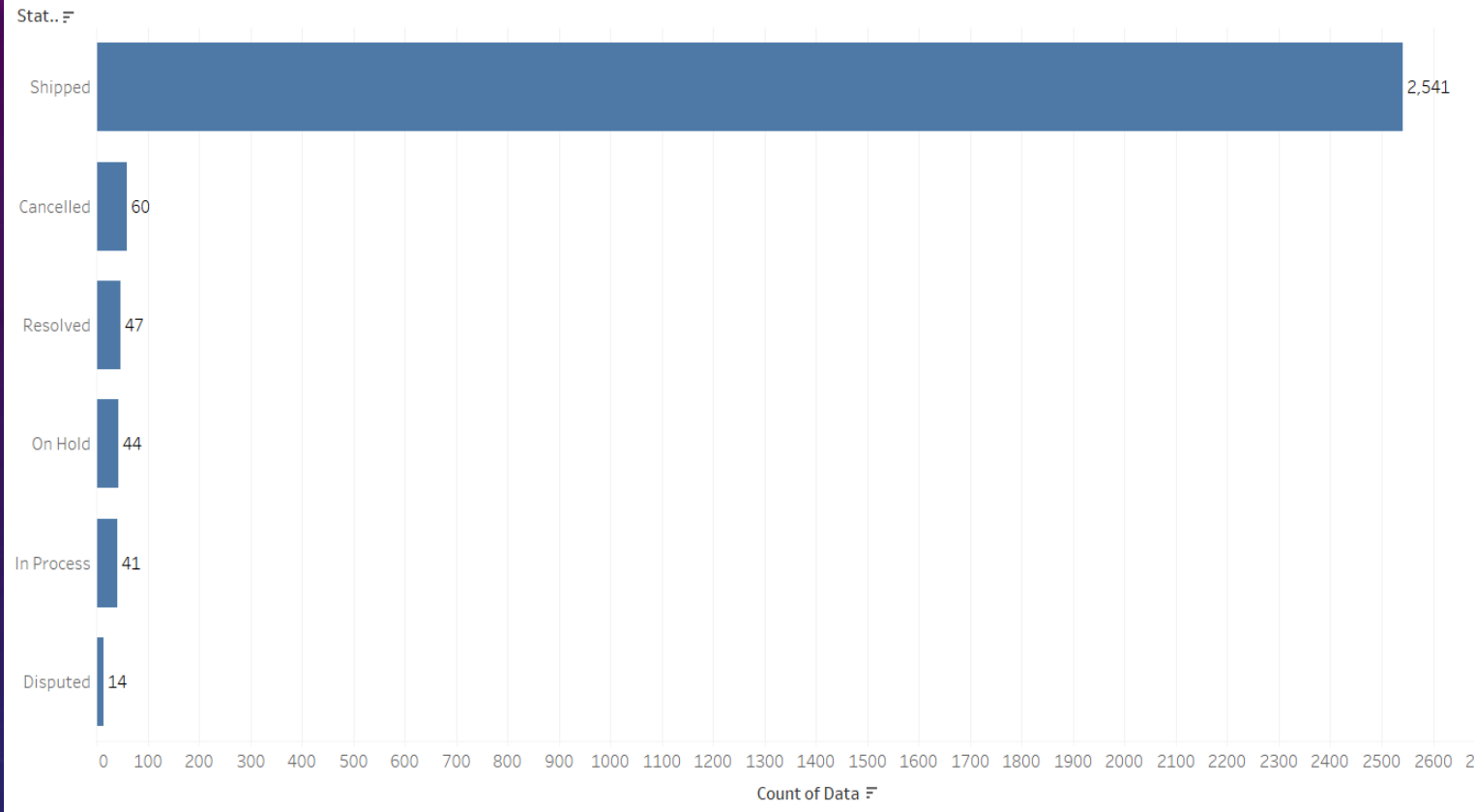
Inference :

- The significant count of 2,541 successfully shipped orders reflects the company's outstanding order fulfilment capabilities and commitment to delivering orders promptly.

SALES ACROSS DEAL-SIZE

Inference :

1. Medium deal-size having highest sales of 5,931,231
2. Small deal-size gets the second highest sales of 2,570,034
3. The concern is all about the Large deal-size which is the lowest in term of sales.
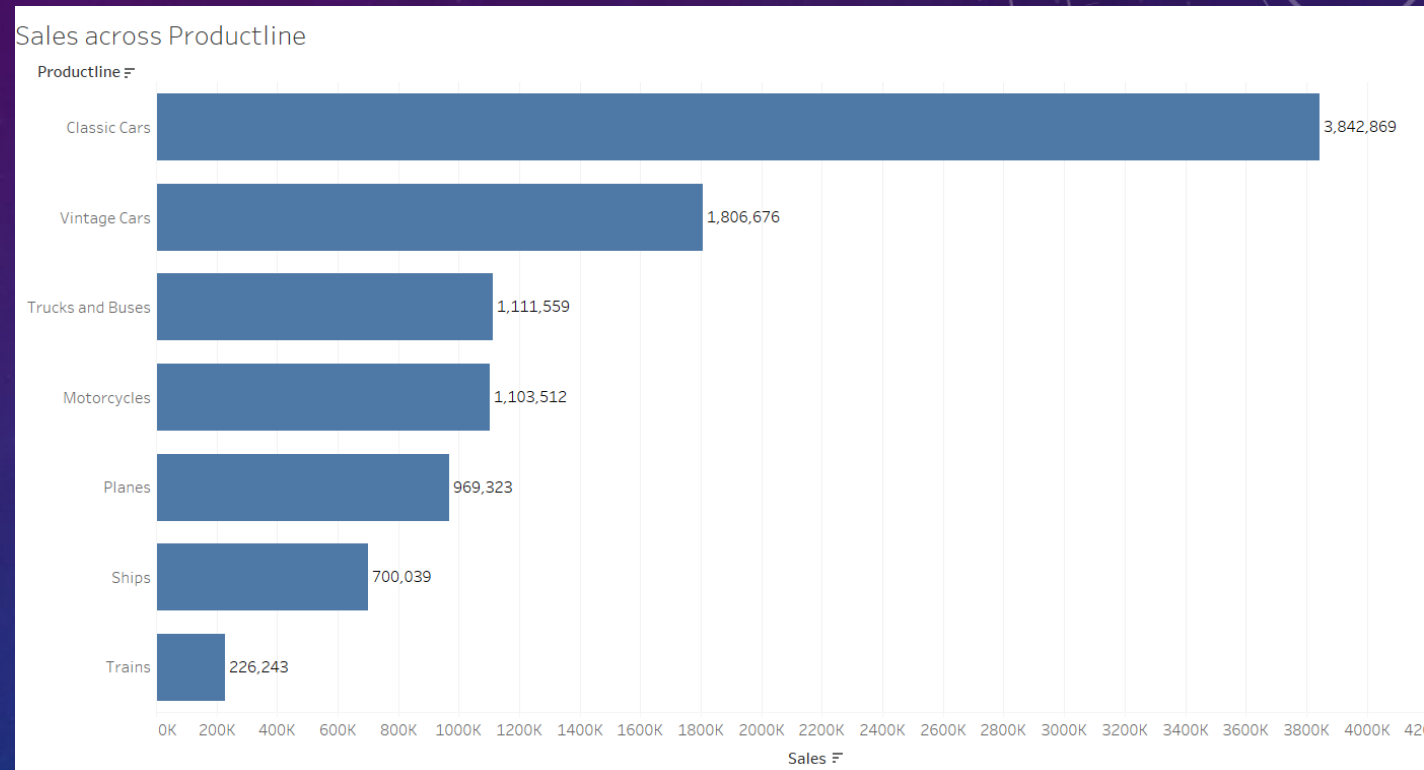

Bivariate-Dealsize vs Sales
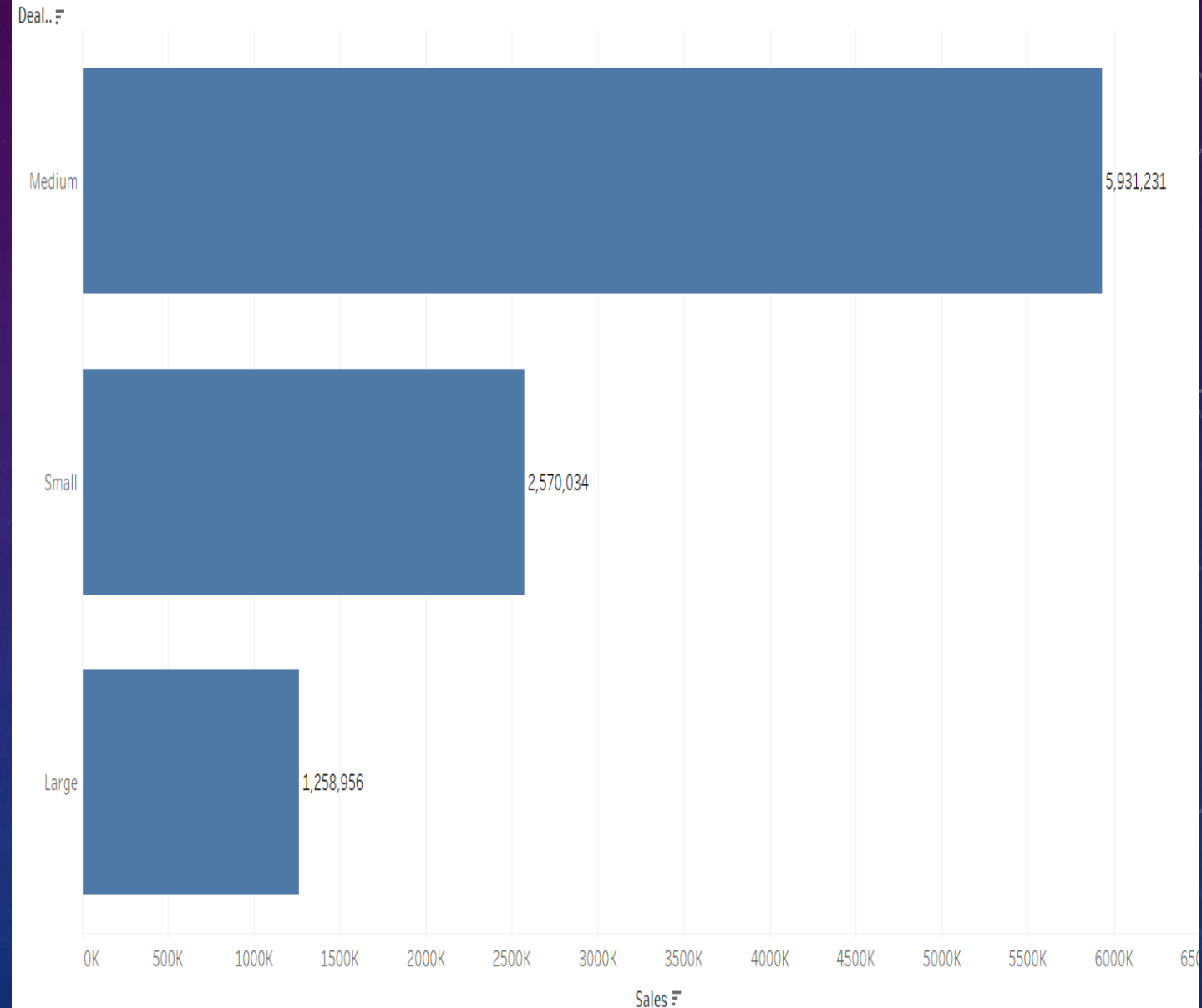
SALES ACROSS PRODUCT LINE

Inference :

1. Classic Cars has the highest sale of 3,842,869 whereas the Trains has the lowest sale of 226,243.
2. Vintage Cars also provides the good number of sales.
3. Company should focus on Trucks, Buses and Motorcycles. As they could gain some more sales in these product lines.
4. Trucks, Buses and Motorcycles should be the target of the company to increase the sales.

Sales across Productline

Productline

Productline	Sales
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

# SALES OVER YEARS

Inference :

1. Sales showed a positive trend and grew significantly from 2018 (3,353,014) to 2019 (4,669,925), suggesting that the business performed well and experienced an increase in customer demand.
2. However, since we only have sales data until May for 2020 (1,737,283), it is difficult to make conclusive statements about the entire year.
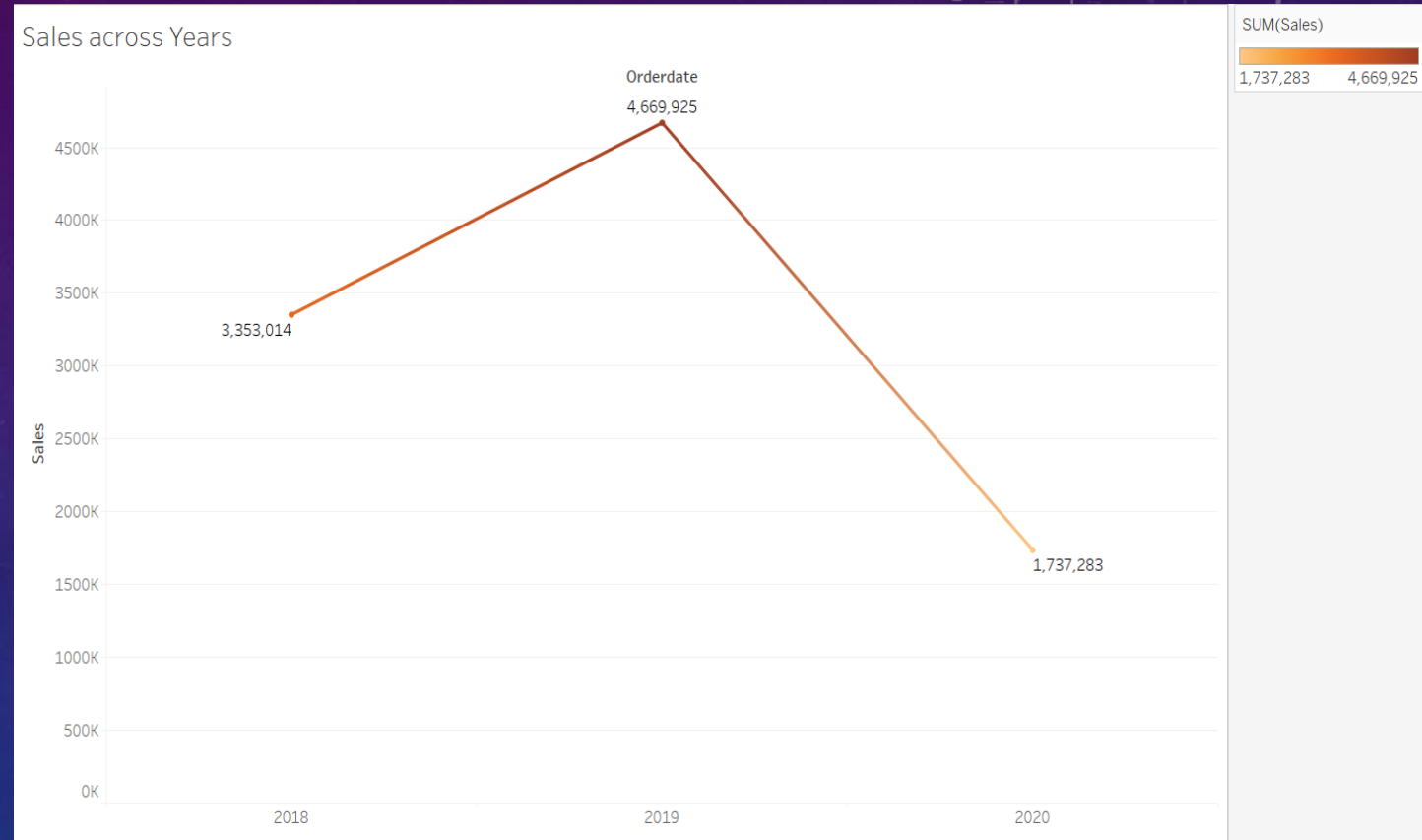
SALES OVER MONTHS

Inference :

1. Sales data shows fluctuations throughout the year.
2. Peak sales occur in November (2,088,537) and October (1,001,377), indicating heightened customer activity likely influenced by seasonal factors.
3. In contrast, June records the lowest sales (454,757), suggesting a potential decrease in customer demand during that period.

Sales over Months
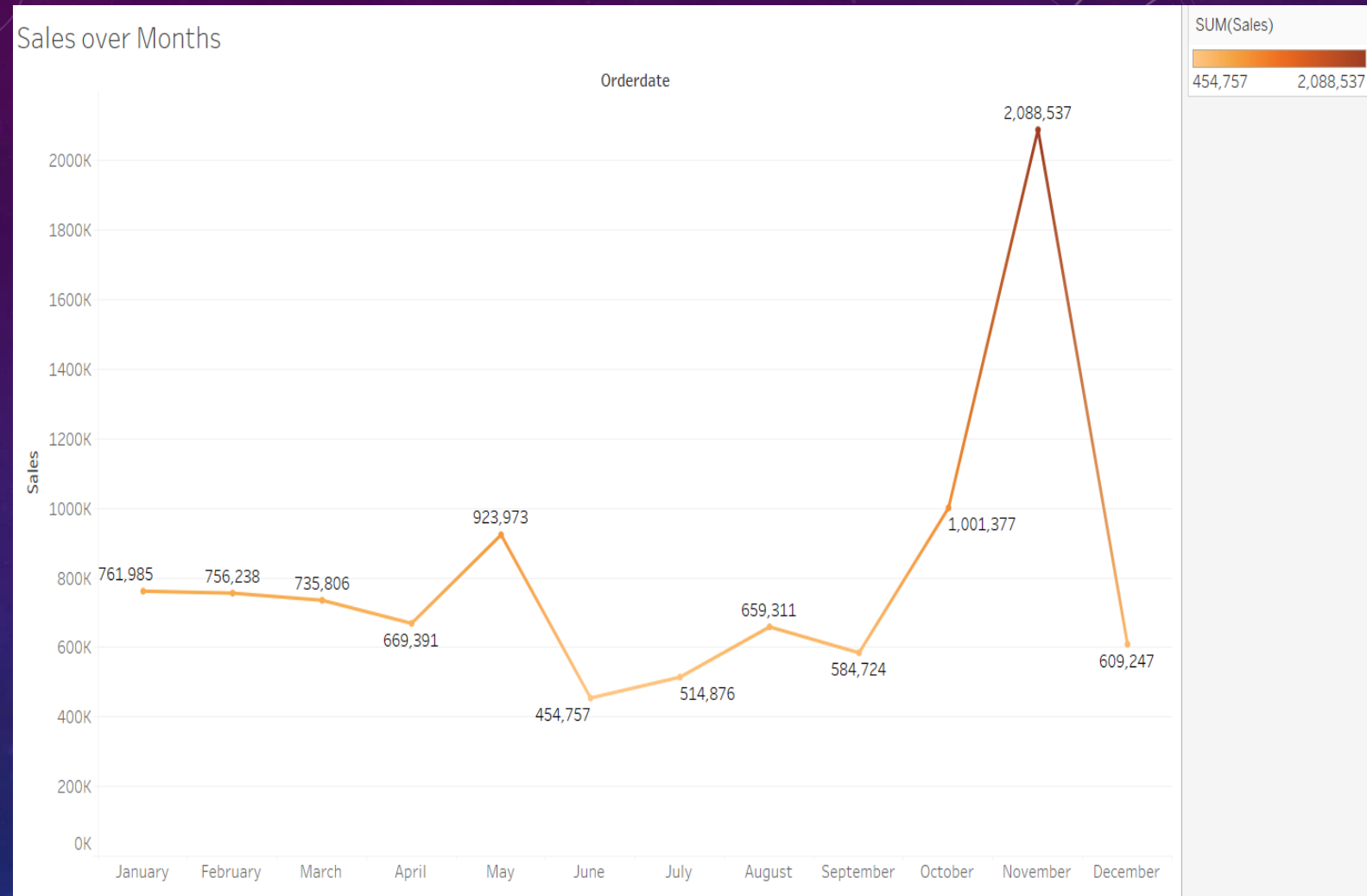
# SALES ACROSS QUARTERS

Inference :

1. Quarterly sales data over three years shows a consistent pattern.
2. Q4 (3,699,161) consistently has higher sales compared to other quarters, indicating increased customer spending towards the end of the year.
3. Q1 (2,254,029) and Q2 (2,048,120) show stable sales figures.
4. Q3 (1,758,911) records the lowest sales among the quarters.

Sales across Quarters
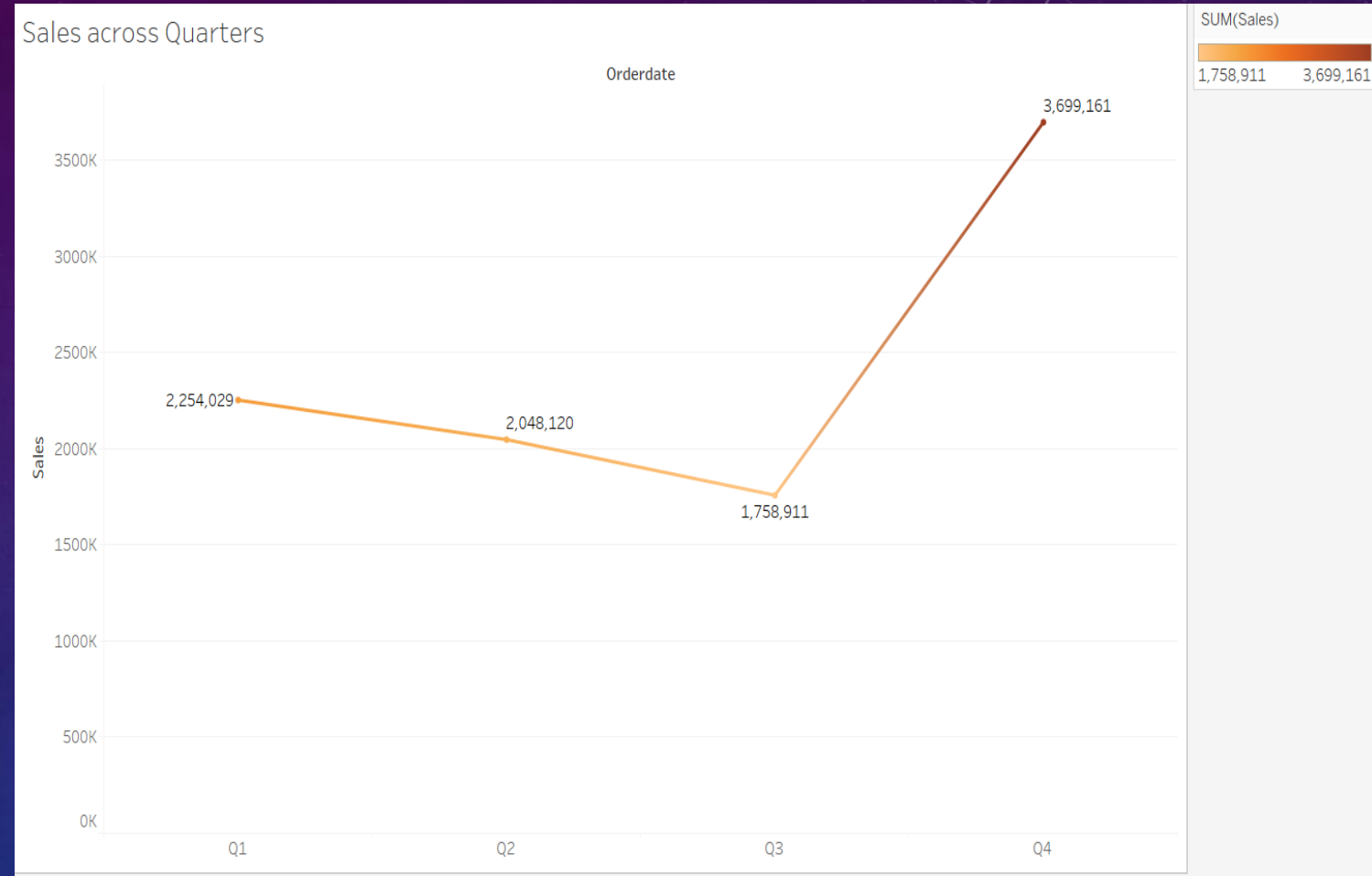
SALES OVER WEEKS

Inference :

1. Throughout the three-year period, Sundays consistently recorded the highest sales, totaling 2,175,121.
2. This trend suggests that customers are more active and inclined to shop during weekends, possibly due to increased leisure time or preference for weekend shopping.
3. In contrast, Wednesdays stand out with the lowest sales among weekdays, totaling 1,374,952.
4. This lower sales figure on Wednesdays indicates that customers may be less engaged or influenced by factors that affect their buying behavior specifically on this day.
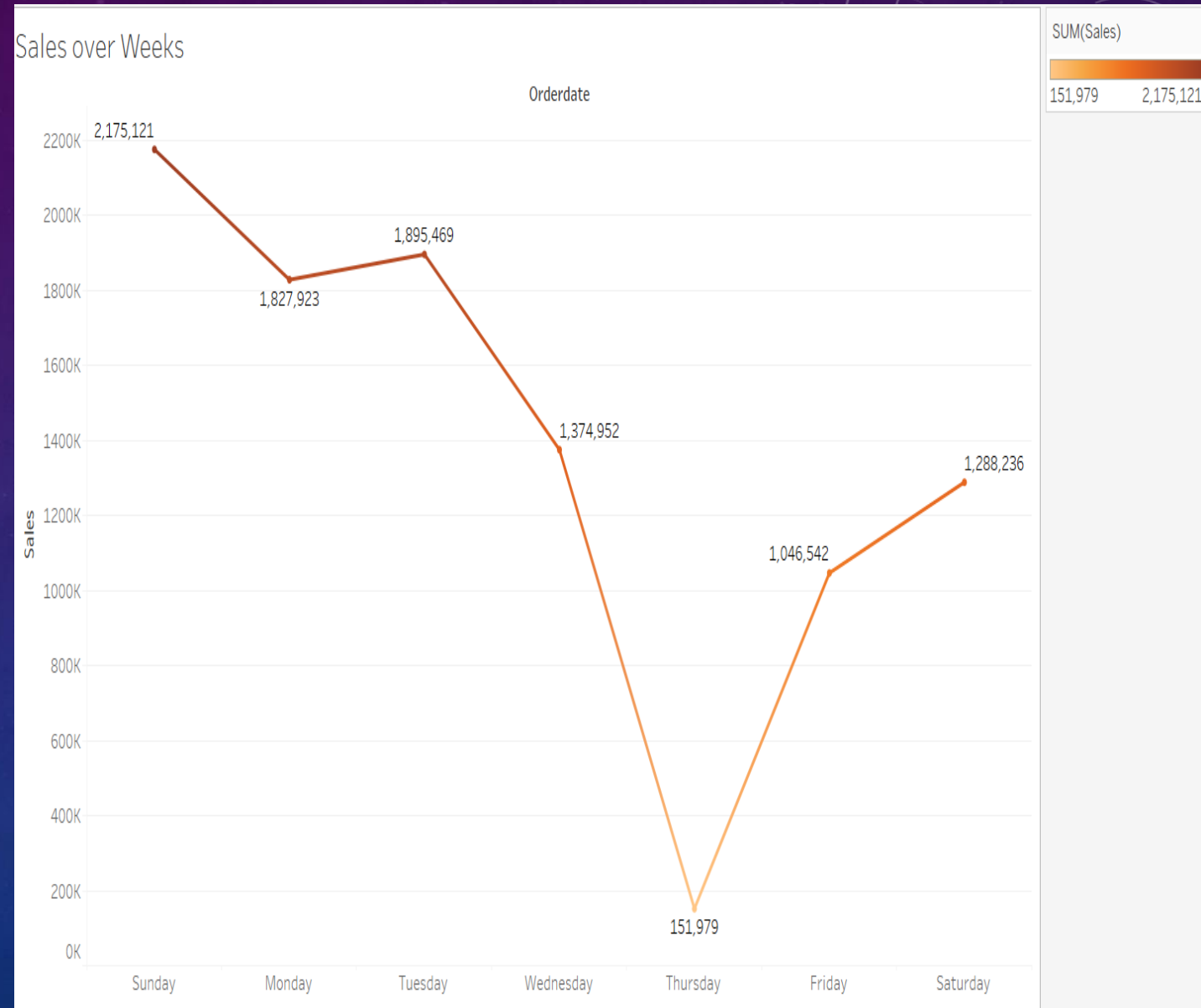
SALES ACROSS COUNTRIES

Inference :

1. Sales data by country reveals significant variations.
2. The USA leads with considerably higher sales (3,355,576), indicating a strong market presence and potentially a larger customer base.
3. Canada (224,079), Singapore (288,488), Denmark (245,637), and Finland (329,582) also show notable sales figures, highlighting significant market opportunities and customer demand in these countries.
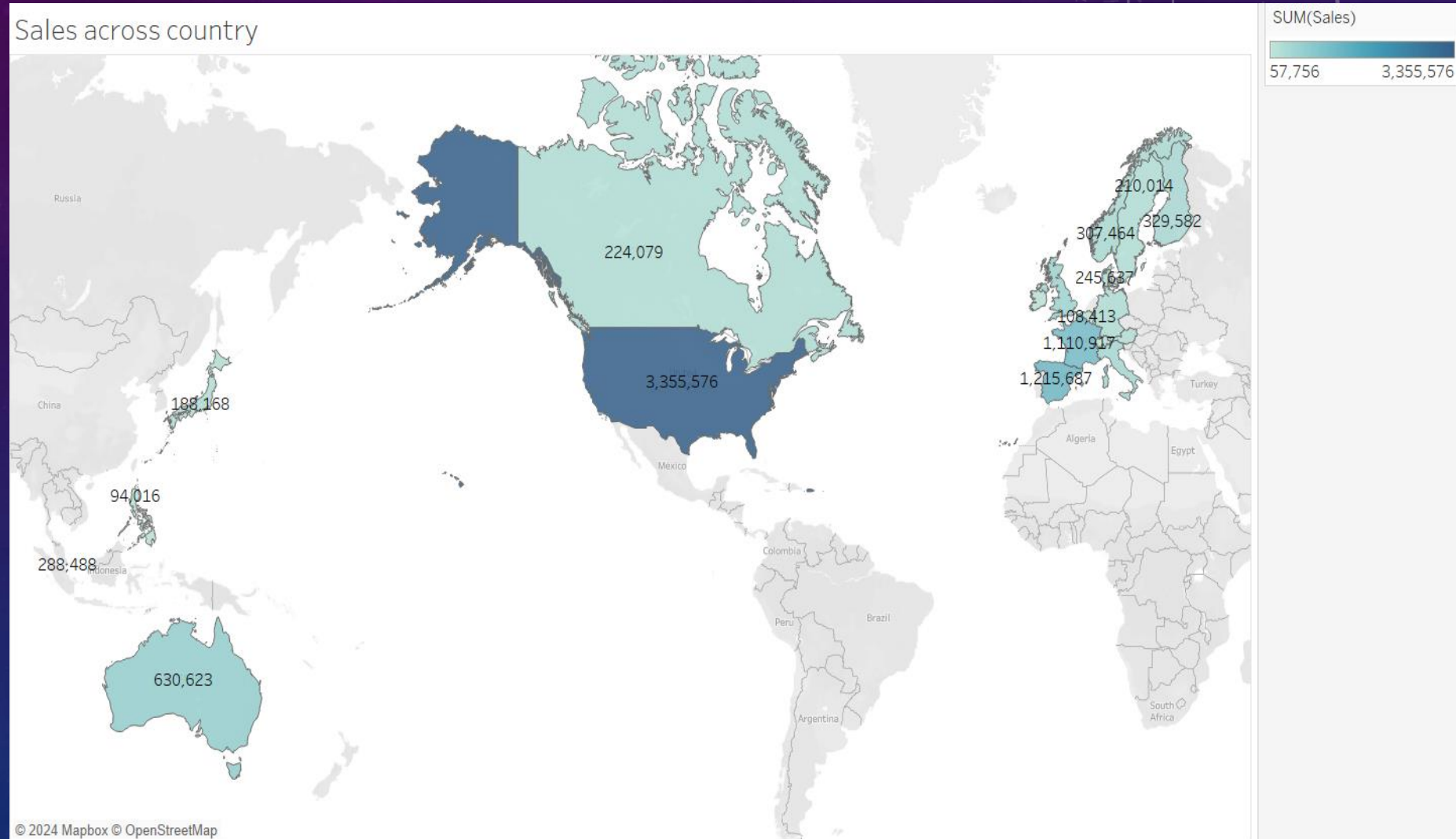
SALES OF PRODUCT LINE ACROSS DEAL SIZE

Inference :

1. Graph analysis shows Classic Cars and Vintage Cars as the most frequently ordered product lines.
2. Motorcycles, Planes, Trucks & Buses, and Ships follow in terms of order frequency.
3. Trains had the lowest number of orders among the product lines.
4. Notably, most orders across all product lines were in the medium and small deal size categories.
5. Higher deal sizes were less common among the popular product lines.

Sales of Productline across dealsize
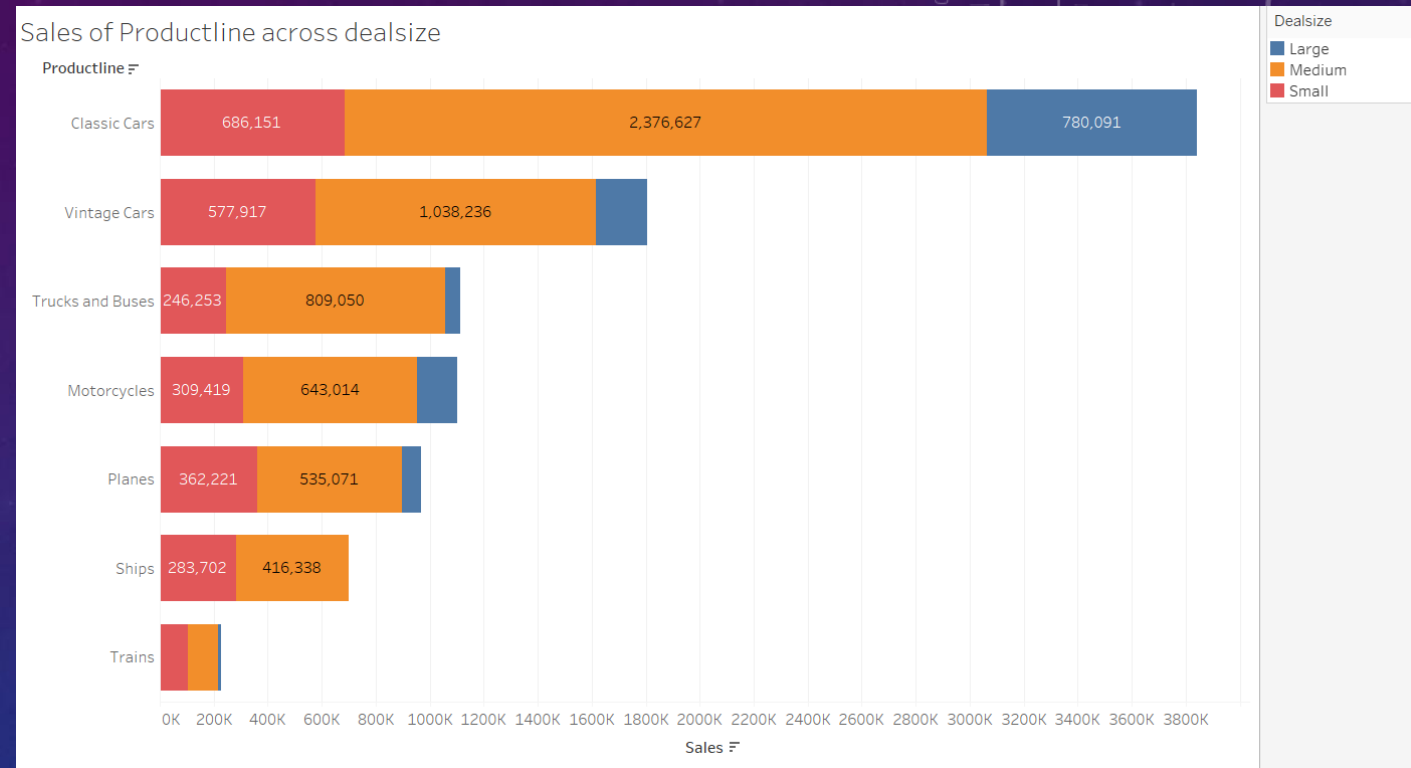
CUSTOMERS SALES ACROSS THE YEARS

Inference :

1. Euro Shopping Channel achieved the top sales position throughout the three-year period.
2. The company consistently maintained its sales performance without significant fluctuations.
3. Mini Gift Distributor Ltd secured the second-highest sales position.
4. It showcased steady and consistent sales growth over the years.
5. Australian Collectors Co ranked third in terms of sales.
6. The company did not record any sales in the year 2020, affecting its overall sales performance.
7. Muscle Machine Inc held the fourth position in terms of sales.
8. Similar to Australian Collectors Co, it also did not generate any sales in the year 2020.

Customer Sales across Years
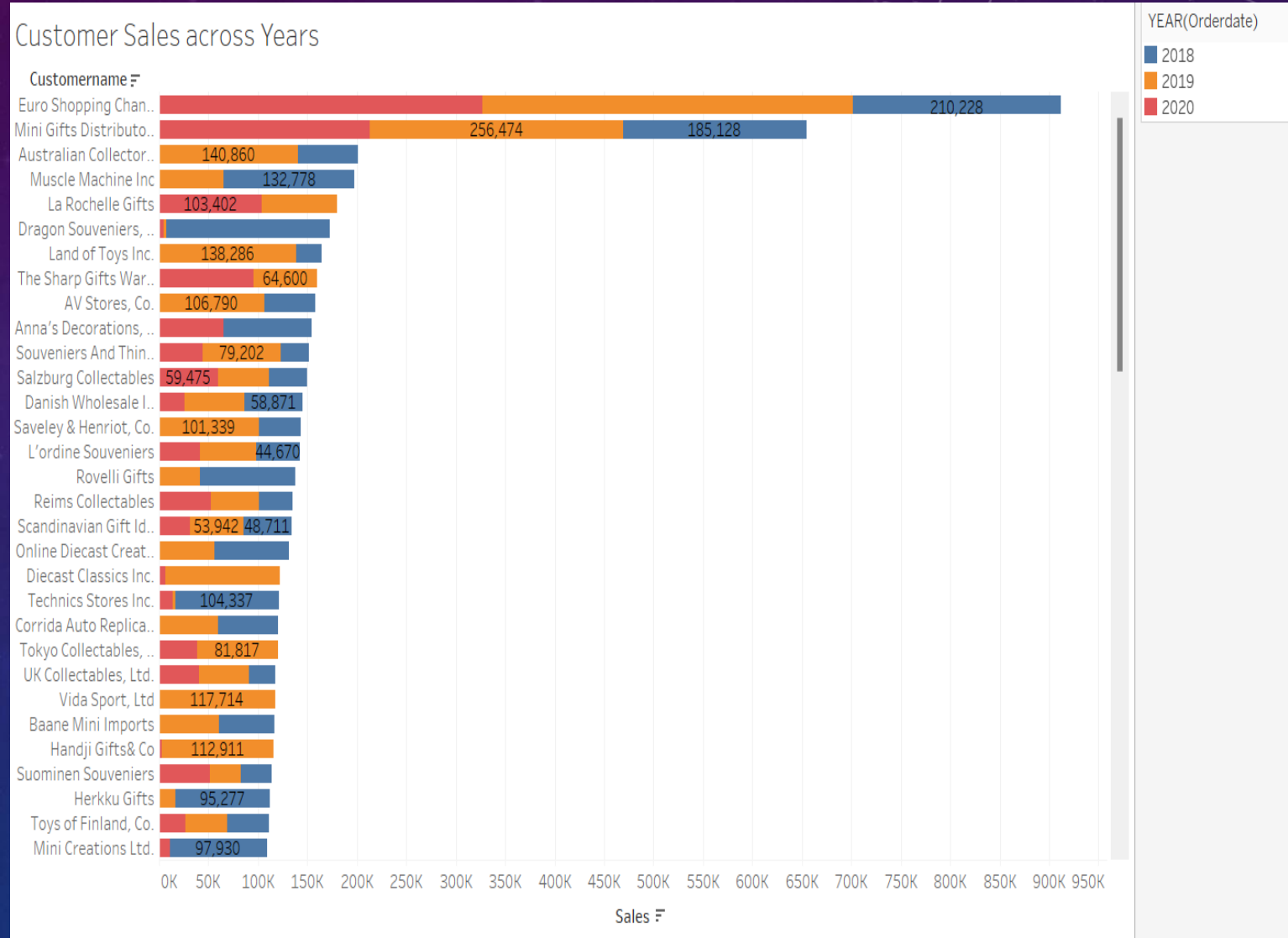
SUMMARY

Product Line Analysis:

- Classic cars and vintage cars stand out as the most popular product lines, indicating a robust demand for these specific vehicle models.

- In contrast, product lines such as ships, trucks, buses, and trains have lower counts, suggesting a specialized customer base or a niche market for these items.

Deal Size Analysis:

- The majority of deals fall into the small and medium categories, with 1,246 and 1,349 counts, respectively.

- This trend indicates that the company primarily engages in transactions of moderate to smaller deal sizes rather than large-scale deals.

Order Status Analysis:

- A significant count of 2,541 successfully shipped orders showcases the company's exceptional order fulfillment capabilities.

- This high number reflects the company's commitment to delivering orders promptly and efficiently.

Quality Ordered Analysis:

- Bins with quality orders of 21, 28, 35, and 42 have the highest counts, indicating their popularity among customers.

- With counts ranging from 576 to 631, these quantities effectively meet the demands of a substantial portion of the customer base.

Sales Analysis:

- Sales exhibited a positive trend, experiencing substantial growth from 2018 (3,353,014) to 2019 (4,669,925).

- This growth suggests that the business performed well and witnessed increased customer demand.

- However, the absence of complete 2020 sales data makes it challenging to draw definitive conclusions about the entire year's performance.

3. CUSTOMER SEGMENTATION USING RFM ANALYSIS

- What is RFM?

- What all parameters used, and assumptions made?

- Showcase the KNIME workflow image.

- What results are there in the output table head?

WHAT IS RFM ?

RFM analysis is a powerful customer segmentation technique widely utilized in marketing and retail industries. It involves evaluating and categorizing customers based on three key metrics:

1.Recency: This metric assesses how recently a customer has made a purchase. Customers who have made recent purchases are often more engaged and responsive to marketing efforts.

2.Frequency: Frequency measures how often a customer makes purchases within a specific timeframe. Customers who make frequent purchases are typically loyal and valuable to the business.

3.Monetary Value: Also known as "Monetary," this metric calculates the total amount of money a customer has spent on purchases. Customers with higher monetary value are considered more profitable and contribute significantly to the company's revenue.

By analyzing these three metrics together, RFM analysis helps businesses identify and understand different customer segments, such as high-value loyal customers, potential churn risks, and new customer opportunities. This information allows companies to tailor their marketing strategies, promotions, and customer experiences to maximize engagement, retention, and profitability.

ASSUMPTIONS

- Greater monetary value implies a more lucrative customer: The belief is that customers who spend more money on purchases are potentially more profitable and valuable for the business.

- Recent purchases serve as a better indicator of customer involvement: The notion is that customers who have recently made purchases are likely more engaged with the company and its products/services, increasing the potential for repeat business or additional sales opportunities.

- More frequent purchases indicate customer loyalty: The assumption is that customers who make purchases more often are more loyal to the company. They tend to have a stronger affinity for the brand, higher satisfaction levels, and are more likely to recommend the company to others.

KNIME WORKFLOW
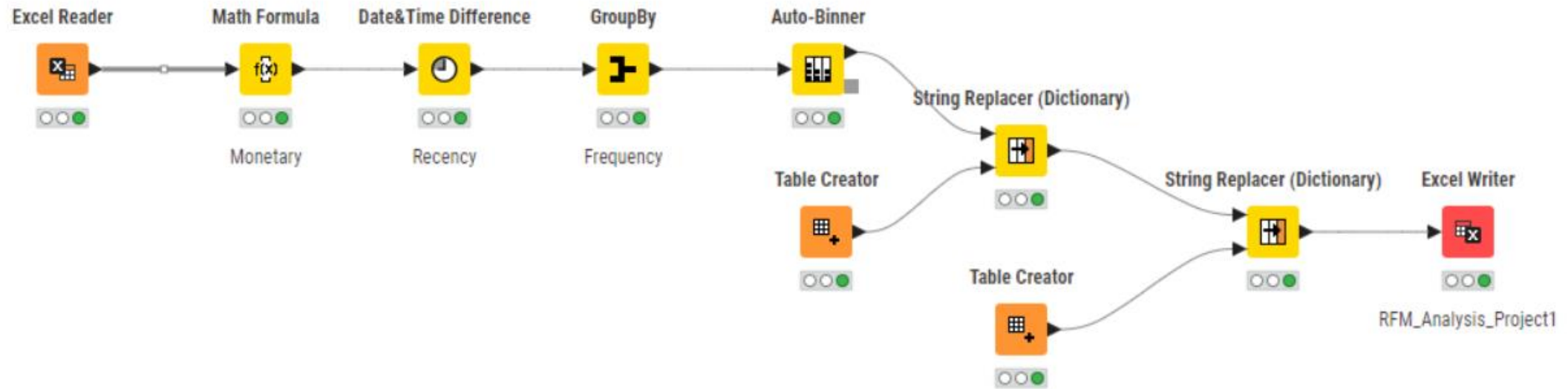
OUTPUT TABLE HEAD

Rows: 89 | Columns: 16

#	RowID	CUSTOM... String	ORDER... Number (d...	QUA... Number...	PRIC... Number ...	ORDER... Local Date	COU... String	MSRP Number (dou...	DEALSIZE String	Monetary Number (dou...	Recency Number (l...	ORD... String	Mon... String	Rec... String	ORDER... String	Monetar... String	Recency [String
1	Row0	AV Stores, Co.	10,253.686	34.863	91.085	2019-10-14	UK	92.843	Medium	157,807.81	1581	Bin 2	Bin 3	Bin 2	M	H	M
2	Row1	Alpha Cognac	10,226.45	34.35	101.16	2018-11-08	France	97.15	Medium	70,488.44	1449	Bin 2	Bin 1	Bin 1	M	L	H
3	Row2	Amica Model...	10,284.5	32.423	110.853	2019-08-17	Italy	107.654	Large	94,117.26	1650	Bin 2	Bin 2	Bin 3	M	M	L
4	Row3	Anna's Decora...	10,250.196	31.935	106.424	2020-03-09	Australia	104.717	Small	153,996.13	1468	Bin 2	Bin 3	Bin 2	M	H	M
5	Row4	Atelier graphi...	10,204.714	38.571	92.239	2019-09-27	France	95.571	Medium	24,179.96	1573	Bin 1	Bin 1	Bin 2	L	L	M
6	Row5	Australian Col...	10,247.522	30.652	90.042	2018-11-21	Australia	88.13	Medium	64,591.46	1407	Bin 2	Bin 1	Bin 1	M	L	H
7	Row6	Australian Col...	10,242.2	35.018	104.59	2019-02-20	Australia	103.527	Medium	200,995.41	1569	Bin 2	Bin 3	Bin 2	M	H	M
8	Row7	Australian Gift...	10,248.133	36.333	110.554	2018-11-06	Australia	111.533	Large	59,469.12	1504	Bin 2	Bin 1	Bin 2	M	L	M
9	Row8	Auto Assoc. &...	10,299.111	35.389	99.488	2019-10-11	France	100.389	Large	64,834.32	1618	Bin 3	Bin 1	Bin 3	H	L	L
10	Row9	Auto Canal Pe...	10,245.889	37.074	94.255	2019-01-15	France	94.852	Medium	93,170.66	1439	Bin 2	Bin 2	Bin 1	M	M	H
11	Row...	Auto/Moto Cl...	10,281	35.875	92.8	2018-06-16	USA	87.375	Medium	26,479.26	1565	Bin 2	Bin 1	Bin 2	M	L	M
12	Row...	Baane Mini Im...	10,205.781	33.812	108.574	2019-10-15	Norway	107.469	Medium	116,599.19	1593	Bin 1	Bin 2	Bin 2	L	M	M
13	Row...	Bavarian Colle...	10,296	28.643	84.289	2019-09-15	Germany	82.714	Medium	34,993.92	1644	Bin 3	Bin 1	Bin 3	H	L	L
14	Row...	Blauer See Au...	10,240.455	36.864	108.031	2018-10-04	Germany	105.818	Medium	85,171.59	1593	Bin 2	Bin 2	Bin 2	M	M	M
15	Row...	Boards & Toys...	10,228	34	89.807	2020-02-08	USA	92.333	Medium	9,129.35	1498	Bin 2	Bin 1	Bin 2	M	L	M
16	Row...	CAF Imports	10,185.308	36	104.963	2019-03-19	Spain	106.923	Large	49,642.05	1824	Bin 1	Bin 1	Bin 3	L	L	L
17	Row...	Cambridge Co...	10,237.545	32.455	101.329	2019-03-10	USA	97.364	Medium	36,163.62	1774	Bin 2	Bin 1	Bin 3	M	L	L
18	Row...	Canadian Gift ...	10,259.5	31.955	105.341	2018-12-05	Canada	106.409	Large	75,238.92	1607	Bin 2	Bin 2	Bin 2	M	M	M
19	Row...	Classic Gift Id...	10,236.143	31.81	103.32	2018-11-13	USA	102.476	Medium	67,506.97	1615	Bin 2	Bin 1	Bin 2	M	L	M
20	Row...	Classic Legen...	10,229.3	36	109.803	2018-10-20	USA	106.65	Medium	77,795.2	1577	Bin 2	Bin 2	Bin 2	M	M	M
21	Row...	Clover Collect...	10,253.688	30.625	112.87	2019-02-12	Ireland	106.875	Large	57,756.43	1643	Bin 2	Bin 1	Bin 3	M	L	L
22	Row...	Collectable Mi...	10,223.12	38.16	91.535	2019-02-19	USA	93.12	Medium	87,489.23	1845	Bin 2	Bin 2	Bin 3	M	M	L
23	Row...	Collectables F...	10,247.458	33.125	97.237	2020-01-20	USA	99.458	Medium	81,577.98	1517	Bin 2	Bin 2	Bin 2	M	M	M
24	Row...	Corrida Auto ...	10,200.75	36.344	105.175	2018-05-28	Spain	102.625	Large	120,615.28	1597	Bin 1	Bin 3	Bin 2	L	H	M
25	Row...	Cruz & Sons C...	10,162.923	36.962	96.08	2018-03-03	Philippin...	97.962	Medium	94,015.73	1582	Bin 1	Bin 2	Bin 2	L	M	M
26	Row...	Daedalus Desi...	10,193.2	34.95	95.474	2018-11-11	France	94.5	Small	69,052.41	1850	Bin 1	Bin 1	Bin 3	L	L	L
27	Row...	Danish Whole...	10,217.361	36.528	108.038	2018-02-11	Denmark	106.417	Large	145,041.6	1431	Bin 1	Bin 3	Bin 1	L	H	H
28	Row...	Diecast Classi...	10,299.097	35.839	108.566	2019-11-02	USA	106.581	Medium	122,138.14	1386	Bin 3	Bin 3	Bin 1	H	H	H

4. INFERENCES FROM RFM ANALYSIS AND IDENTIFIED SEGMENTS

- Who are your best customers?

- Which customers are on the verge of churning?

- Who are your lost customers?

- Who are your loyal customers?

TOP 5 BEST CUSTOMERS

CUSTOMERNAME	ORDERNUMBER [Binned](H,M,L)	Monetary [Binned](H,M,L)	Recency [Binned](H,M,L)
Diecast Classics Inc.	H	H	H
Euro Shopping Channel	H	H	H
L'ordine Souveniers	H	H	H
La Rochelle Gifts	H	H	H
Reims Collectables	H	H	H

H = High
M = Moderate
L = Low

Customers with High Recency, High Monetary and High Frequency falls under the best customers category having RFM score is greater than others.

CUSTOMERS WHO'S ON THE VERGE OF CHURNING

CUSTOMERNAME	ORDERNUMBER [Binned](H,M,L)	Monetary [Binned](H,M,L)	Recency [Binned](H,M,L)
Amica Models & Co.	M	M	L
Collectable Mini Designs Co.	M	M	L
Marta's Replicas Co.	H	M	L
Norway Gifts By Mail, Co.	H	M	L
Royal Canadian Collectables, Ltd.	M	M	L
Super Scale Inc.	M	M	L
Vida Sport, Ltd	M	M	L

H = High
M = Moderate
L = Low

Customers with low Recency, High and Moderate Monetary and High and Moderate Frequency comes under the verge of churning.

WHO ARE YOUR LOST CUSTOMERS?

CUSTOMERNAME	ORDERNUMBER [Binned](H,M,L)	Monetary [Binned](H,M,L)	Recency [Binned](H,M,L)
CAF Imports	L	L	L
Daedalus Designs Imports	L	L	L
Double Decker Gift Stores, Ltd	L	L	L
Iberia Gift Imports, Corp.	L	L	L
Osaka Souveniers Co.	L	L	L
Signal Collectibles Ltd.	L	L	L
West Coast Collectables Co.	L	L	L

H = High
M = Moderate
L = Low

Customers with Low Recency, Low Monetary and Low Frequency falls under the lost customers category.

WHO ARE YOUR LOYAL CUSTOMERS?

CUSTOMERNAME	ORDERNUMBER [Binned](H,M,L)	Monetary [Binned](H,M,L)	Recency [Binned](H,M,L)
Diecast Classics Inc.	H	H	H
Euro Shopping Channel	H	H	H
L'ordine Souveniers	H	H	H
La Rochelle Gifts	H	H	H
Reims Collectables	H	H	H
Souveniers And Things Co.	H	H	H
The Sharp Gifts Warehouse	H	H	H

Above are the top and Loyal customers. As they has solid order history or the frequency with the bulk number of orders as Monetary and having the most recent interactions with the company as Recency.

5. RECOMMENDATIONS

- <u>Implement Strategies for Customer Retention:</u>

Develop personalized initiatives such as loyalty programs, customized incentives, and exclusive promotions to maintain the engagement of customers who show signs of churning.

- <u>Utilize Continuous Monitoring and Analysis:</u>

Consistently track customer behaviors, sales data, and key performance metrics to detect shifts in patterns and take proactive measures to address potential issues.

- <u>Leverage Upselling and Cross-Selling Opportunities:</u>

Identify chances to upsell or cross-sell to existing customers, particularly those with significant value. Recommend relevant products or services based on their buying history to boost their overall spending and foster stronger loyalty.