

MARKETING AND RETAIL ANALYSIS

AUTOMOBILE PARTS MANUFACTURING COMPANY

**GAUTHAM GONGADA
POST GRADUATE PROGRAM IN
DATA SCIENCE AND BUSINESS
ANALYTICS**

TABLE OF CONTENTS

01 Basic Data Summary

02 Exploratory Data Analysis

03 RFM

04 Inference from RFM

Problem Statement 1 :

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.

OBJECTIVE :

The aim of this analysis was to discern the fundamental purchasing trends among customers, offer insightful details regarding customer behavior, and suggest tailored marketing tactics suited to various customer segments.

Basic Data Summary

- The dataset contains records of 2,747 transactions conducted by customers of an automobile parts manufacturing company.
- It comprises 20 columns, incorporating 2 float, 6 integer, and 12 object data types, offering a wide array of information about both transactions and customers.
- There are no null values.
- There are 4730 duplicate entries within the dataset. On average, customers ordered 35 items per transaction, with a minimum of 6 and a maximum of 97.
- The average duration since the last order is 1,757.09 days, ranging from a minimum of 42 days to 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.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2747 entries, 0 to 2746
Data columns (total 20 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   ORDERNUMBER      2747 non-null   int64  
 1   QUANTITYORDERED 2747 non-null   int64  
 2   PRICEEACH        2747 non-null   float64 
 3   ORDERLINENUMBER  2747 non-null   int64  
 4   SALES            2747 non-null   float64 
 5   ORDERDATE        2747 non-null   int64  
 6   DAYS_SINCE_LASTORDER 2747 non-null   int64  
 7   STATUS            2747 non-null   object  
 8   PRODUCTLINE      2747 non-null   object  
 9   MSRP              2747 non-null   int64  
 10  PRODUCTCODE      2747 non-null   object  
 11  CUSTOMERNAME    2747 non-null   object  
 12  PHONE             2747 non-null   object  
 13  ADDRESSLINE1    2747 non-null   object  
 14  CITY              2747 non-null   object  
 15  POSTALCODE       2747 non-null   object  
 16  COUNTRY           2747 non-null   object  
 17  CONTACTLASTNAME 2747 non-null   object  
 18  CONTACTFIRSTNAME 2747 non-null   object  
 19  DEALSIZE          2747 non-null   object  
dtypes: float64(2), int64(6), object(12)
memory usage: 429.3+ KB
```

Null Values	Column
ORDERNUMBER	0
QUANTITYORDERED	0
PRICEEACH	0
ORDERLINENUMBER	0
SALES	0
ORDERDATE	0
DAYS_SINCE_LASTORDER	0
STATUS	0
PRODUCTLINE	0
MSRP	0
PRODUCTCODE	0
CUSTOMERNAME	0
PHONE	0
ADDRESSLINE1	0
CITY	0
POSTALCODE	0
COUNTRY	0
CONTACTLASTNAME	0
CONTACTFIRSTNAME	0
DEALSIZE	0
dtype:	int64

OrderNumber: This column contains the order number for each order. The minimum order number is 10100, the maximum order number is 10425, and the average order number is 10259.76.

QuantityOrdered: This column contains the quantity ordered for each order. The minimum quantity ordered is 6, the maximum quantity ordered is 97, and the average quantity ordered is 35.10.

PriceEach: This column contains the price of each item ordered. The minimum price is \$26.88, the maximum price is \$252.87, and the average price is \$101.09.

OrderLineNumber: This column contains the order line number for each order. The minimum order line number is 1, the maximum order line number is 18, and the average order line number is 6.49.

	count	mean	std	min	25%	50%	75%	max	
ORDERNUMBER	2747.0	10259.761558	91.877521	10100.00	10181.000	10264.00	10334.500	10425.00	
QUANTITYORDERED	2747.0	35.103021	9.762135	6.00	27.000	35.00	43.000	97.00	
PRICEEACH	2747.0	101.098951	42.042548	26.88	68.745	95.55	127.100	252.87	
ORDERLINENUMBER	2747.0	6.491081	4.230544	1.00	3.000	6.00	9.000	18.00	
SALES	2747.0	3553.047583	1838.953901	482.13	2204.350	3184.80	4503.095	14082.80	
ORDERDATE	2747.0	43598.914088	230.231295	43106.00	43412.000	43640.00	43786.000	43982.00	
DAYS_SINCE_LASTORDER	2747.0	1757.085912	819.280576	42.00	1077.000	1761.00	2436.500	3562.00	
MSRP	2747.0	100.691664	40.114802	33.00	68.000	99.00	124.000	214.00	

Sales: This column contains the sales total for each order. The minimum sales total is \$482.13, the maximum sales total is \$14082.80, and the average sales total is \$3553.05.

OrderDate: This column contains the order date for each order. The minimum order date is 43106 (which corresponds to a date in December 2017), the maximum order date is 43982 (which corresponds to a date in March 2018), and the average order date is 43598.91.

DaysSinceLastOrder: This column contains the number of days since the last order for each order. The minimum number of days since the last order is 42, the maximum number of days since the last order is 3562, and the average number of days since the last order is 1757.08.

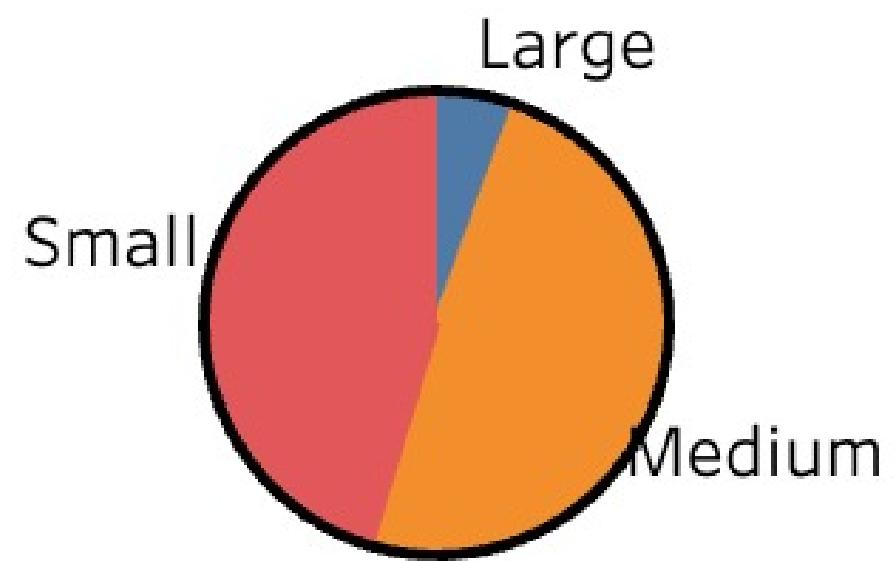
MSRP: This column contains the MSRP (Manufacturer's Suggested Retail Price) for each item ordered. The minimum MSRP is \$33.00, the maximum MSRP is \$214.00, and the average MSRP is \$100.69.

In conclusion, the table shows that the average order consists of approximately 35 items, with a total sales price of \$3553.05. The items are typically priced around \$101.09 each.

Exploratory Data Analysis:

The data reveals that there is a notable discrepancy in customer preferences regarding order sizes. Specifically, while only 152 orders are categorized as "large" deals, a substantially larger number of customers, totaling 1246, express interest in "small" orders. Moreover, the "medium" deal size attracts an even higher number of orders, with 1349 in total.

dealsize



Dealsize

- (All)
- Large
- Medium
- Small

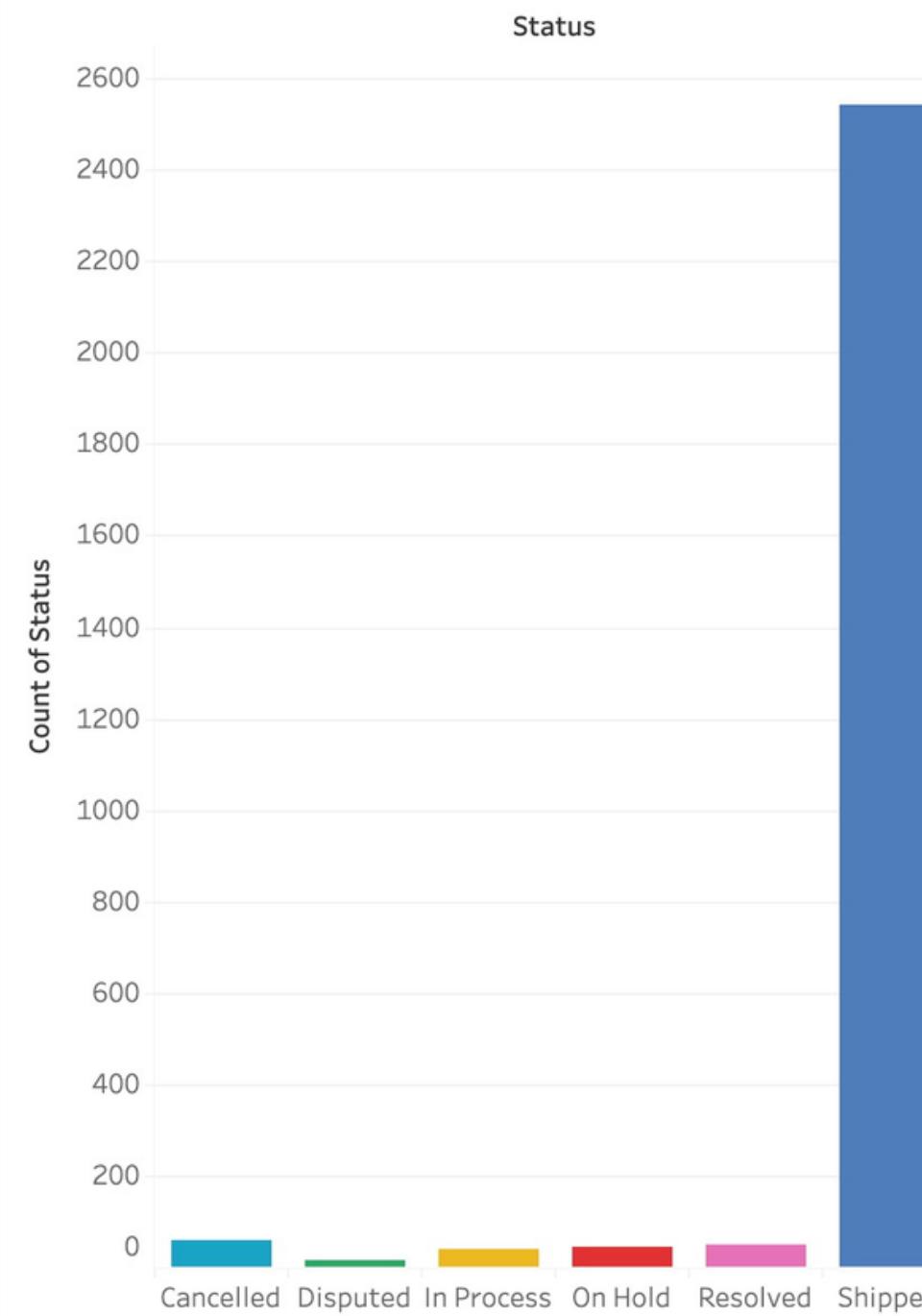
Dealsize

- █ Large
- █ Medium
- █ Small

This suggests that the majority of customers exhibit a preference for smaller or medium-sized orders over larger ones. This insight could prompt businesses to tailor their strategies accordingly, focusing on catering to the demand for smaller and medium-sized deals to better align with customer preferences and maximize sales opportunities.

Order status

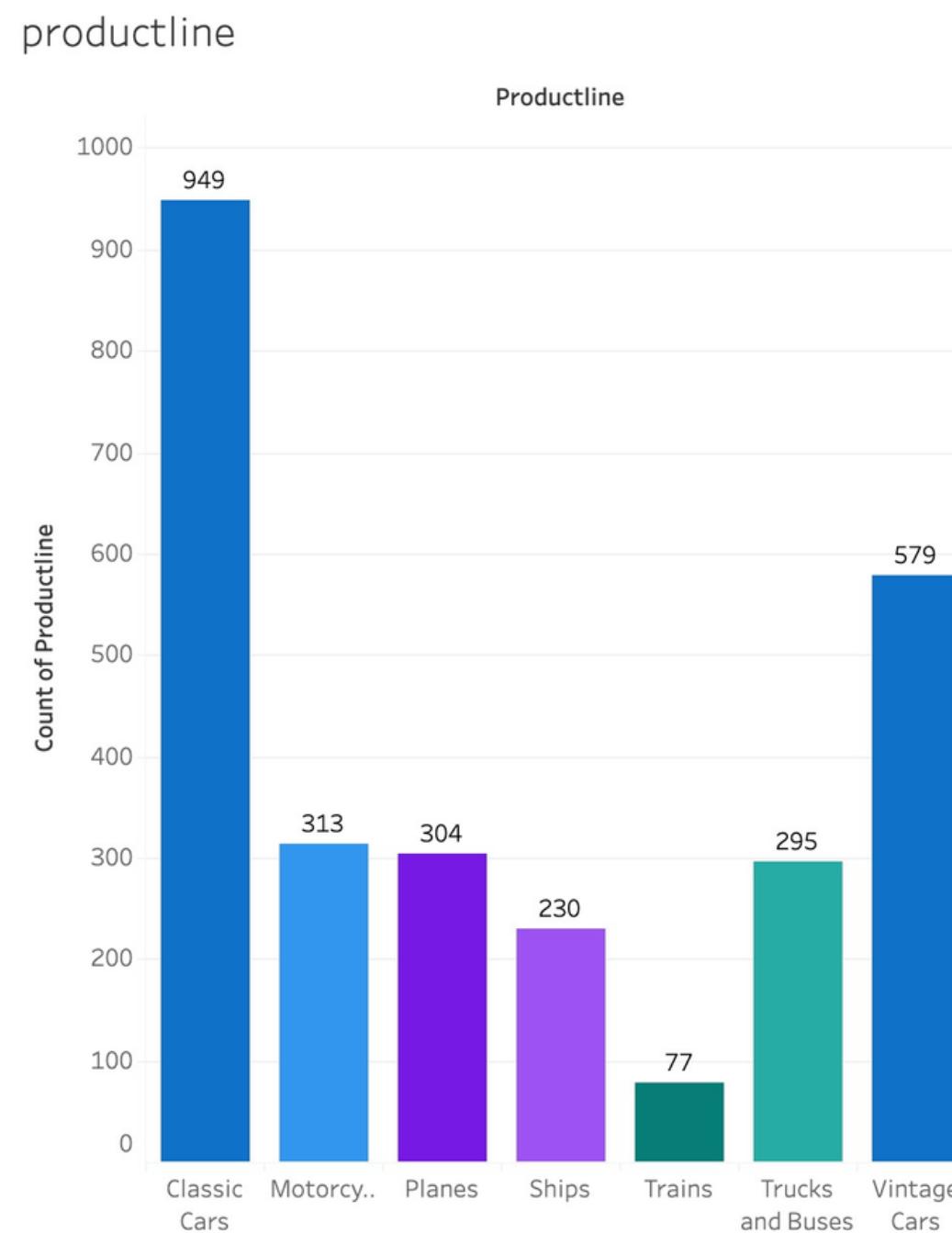
order status



We shipped over **92%** of the 2747 orders, with less than **2.5%** of them being canceled.

Row Labels	Count of STATUS	
Cancelled	60	2.18%
Disputed	14	0.51%
In Process	41	1.49%
On Hold	44	1.60%
Resolved	47	1.71%
Shipped	2541	92.50%
Grand Total	2747	

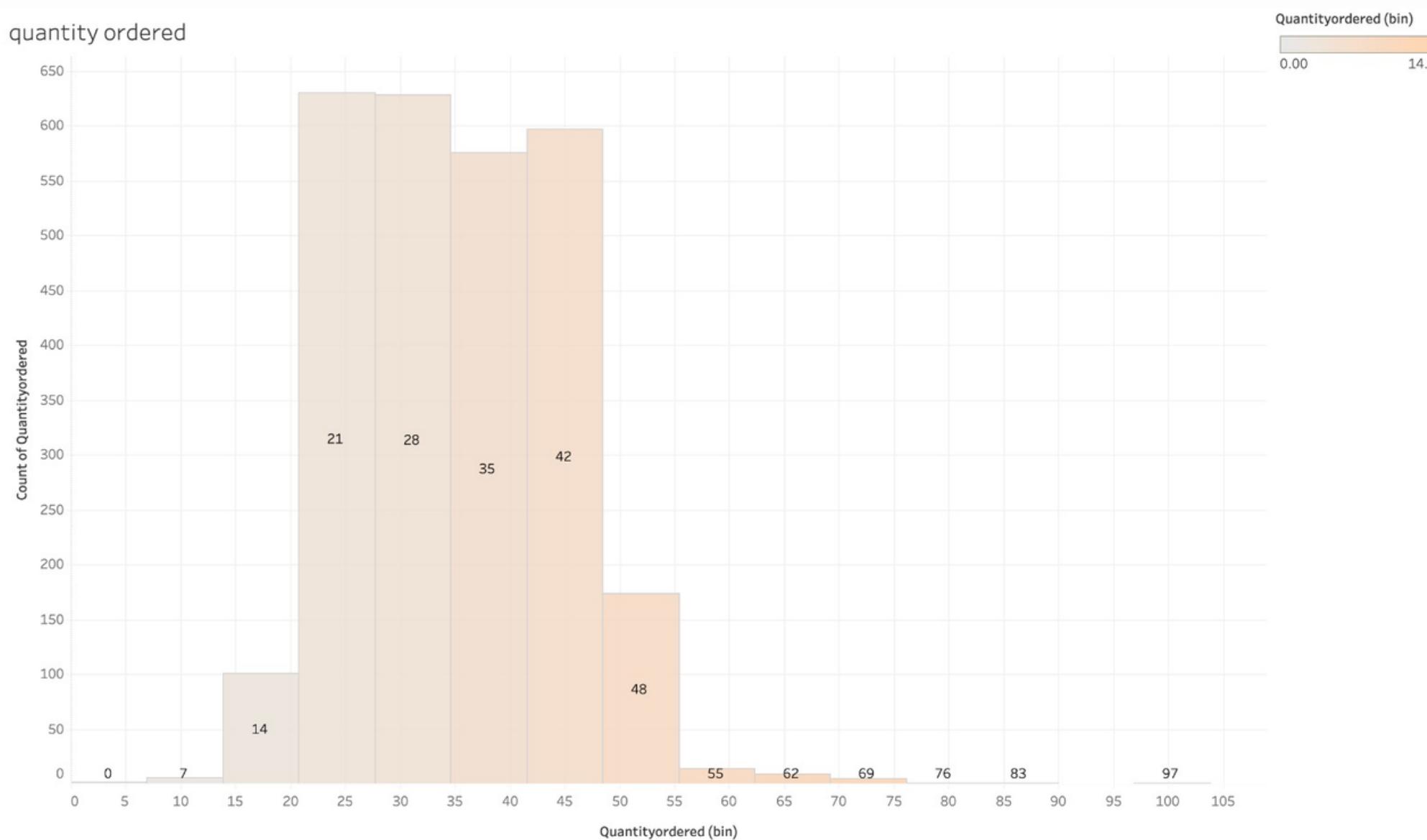
Product Line Status



Classic cars account for the largest percentage of shipped orders (35%), followed by vintage cars (21%). Least of all the categories is trains with only 3%.

Row Labels	Count of PRODUCTLINE	
Classic Cars	949	35%
Motorcycles	313	11%
Planes	304	11%
Ships	230	8%
Trains	77	3%
Trucks and Buses	295	11%
Vintage Cars	579	21%
Grand Total	2747	

Quantity_Ordered



The average of quantity ordered is 35 items per order ,with the highest being 97 Items.

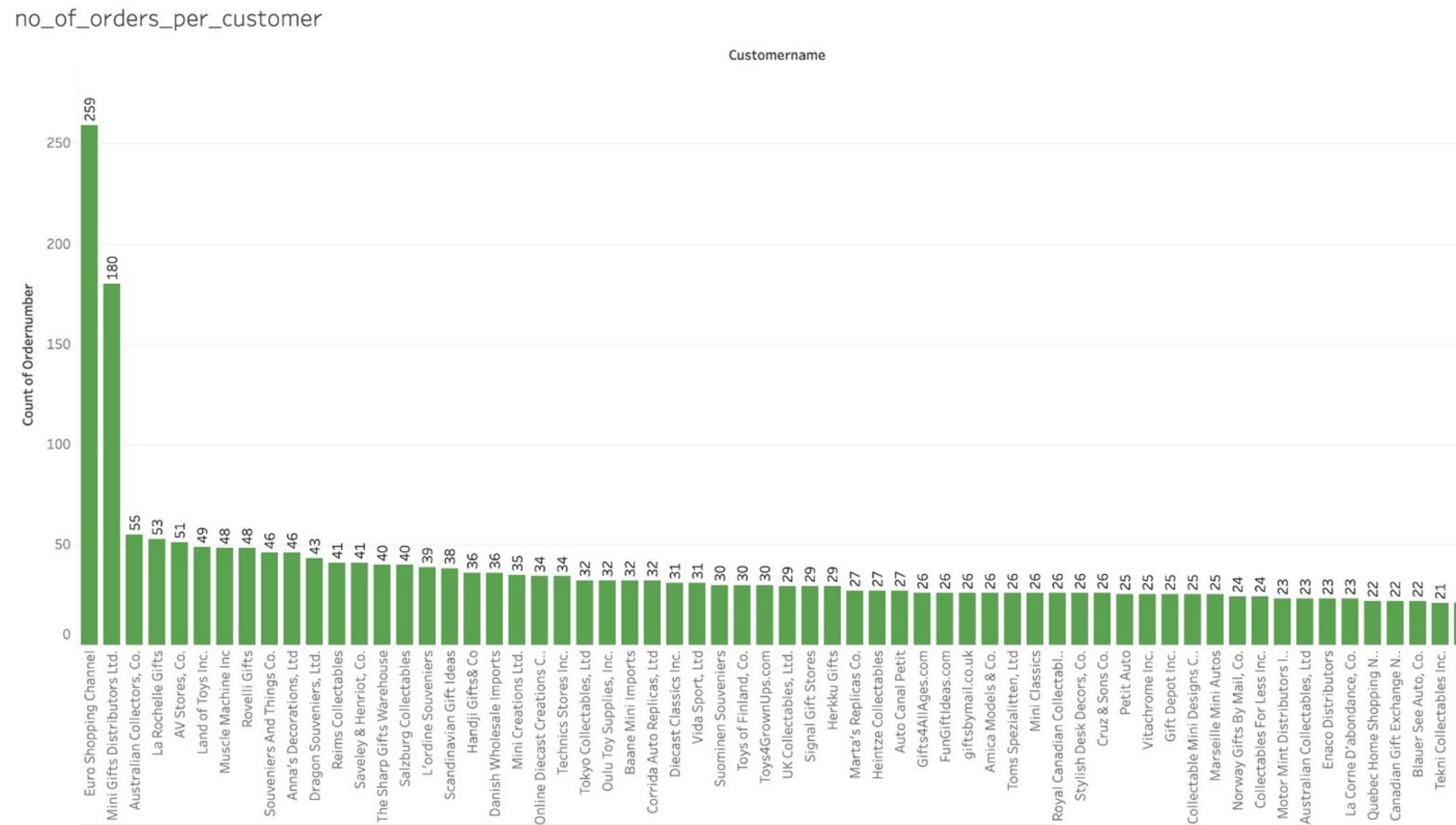
Country data



The highest of the orders are from USA with 34% ,followed by Spain and France with 12% each and the least is from Ireland with less than 1%

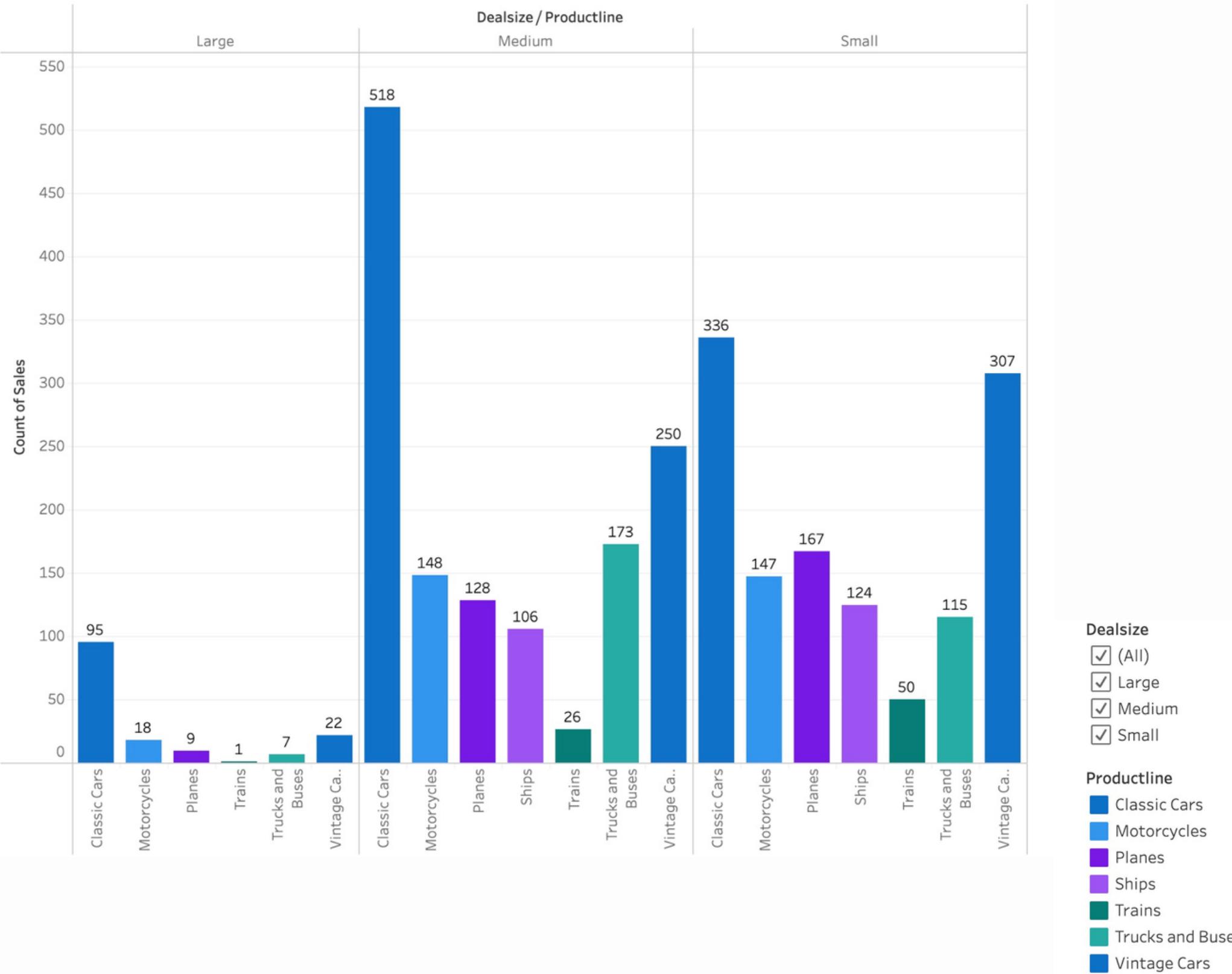
Orders per customer

Euro shopping channel is the customer who placed the highest number of orders contributing to 9.4% followed by Mini Gift Distributors with 7%



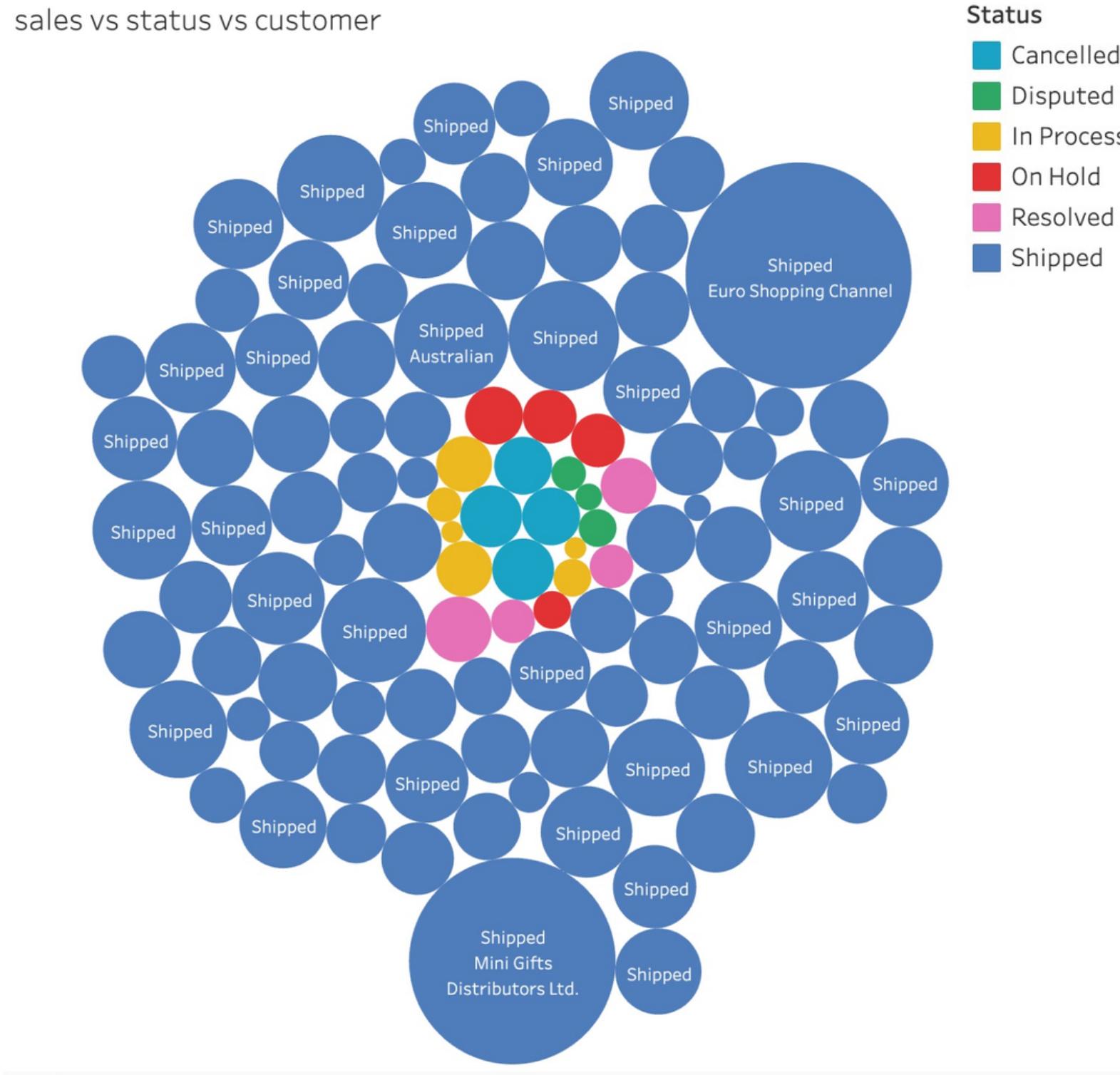
Sales per deal size per product line

sales vs dealsize vs productline



Classic cars have been the highest selling product category for both high and medium sized deals .Whereas for small deal size category ,there is almost equal share between Classic cars and Vintage cars

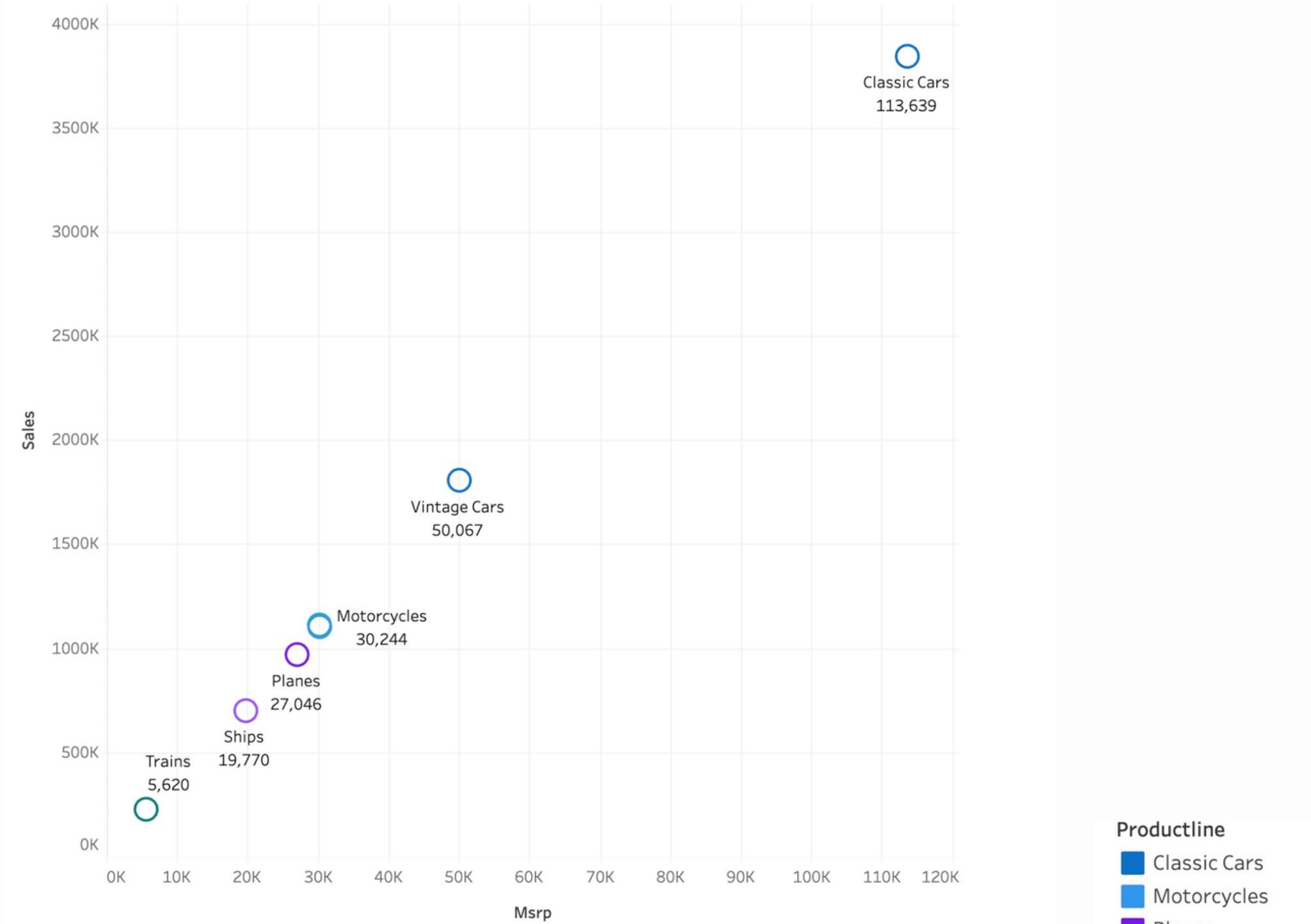
Status per customer



The flowchart represents the order processing status for customers. It outlines various paths an order can follow, including placing an order, which may result in either shipment or non-shipment. Shipped orders can further diverge, being sent to either a Euro Shopping Channel customer or an Australian customer.

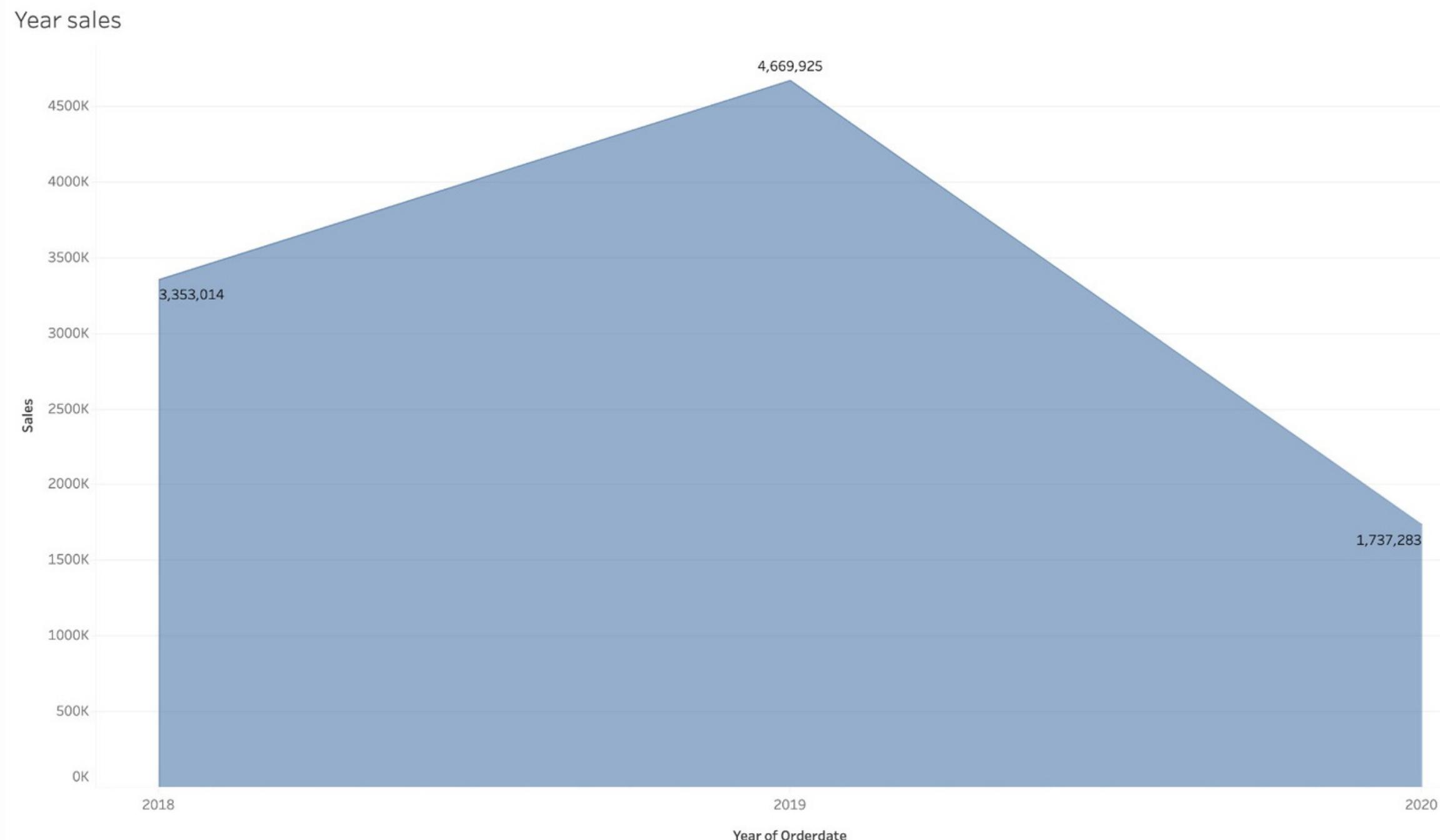
MSRP vs productline

msrp vs productline vs sales



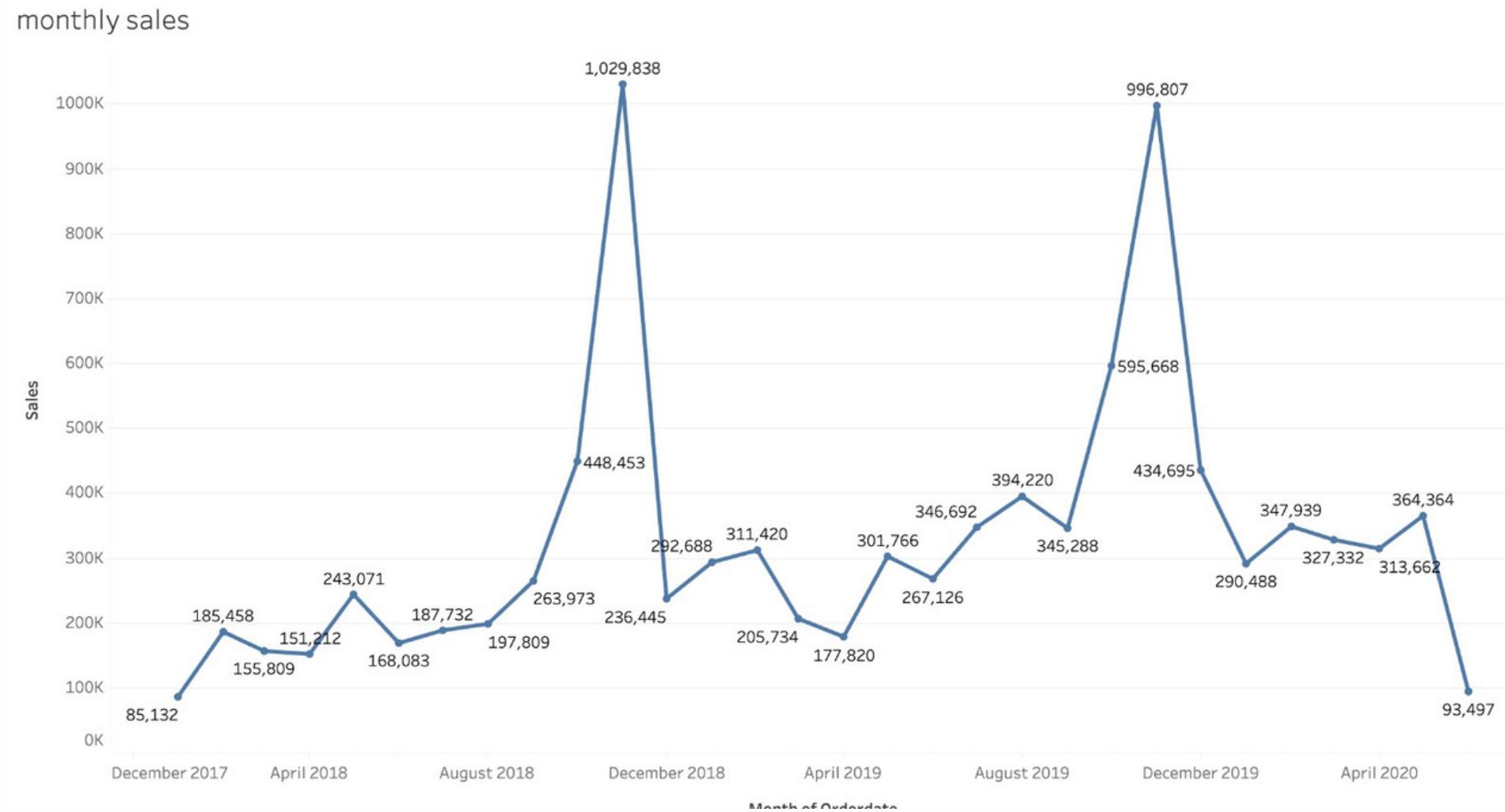
Classic cars category have the highest MSRP contributing to 35% of the total MSRP with 43% contribution to sales with the least contribution from trains category at less than 3%

Yearly Sales



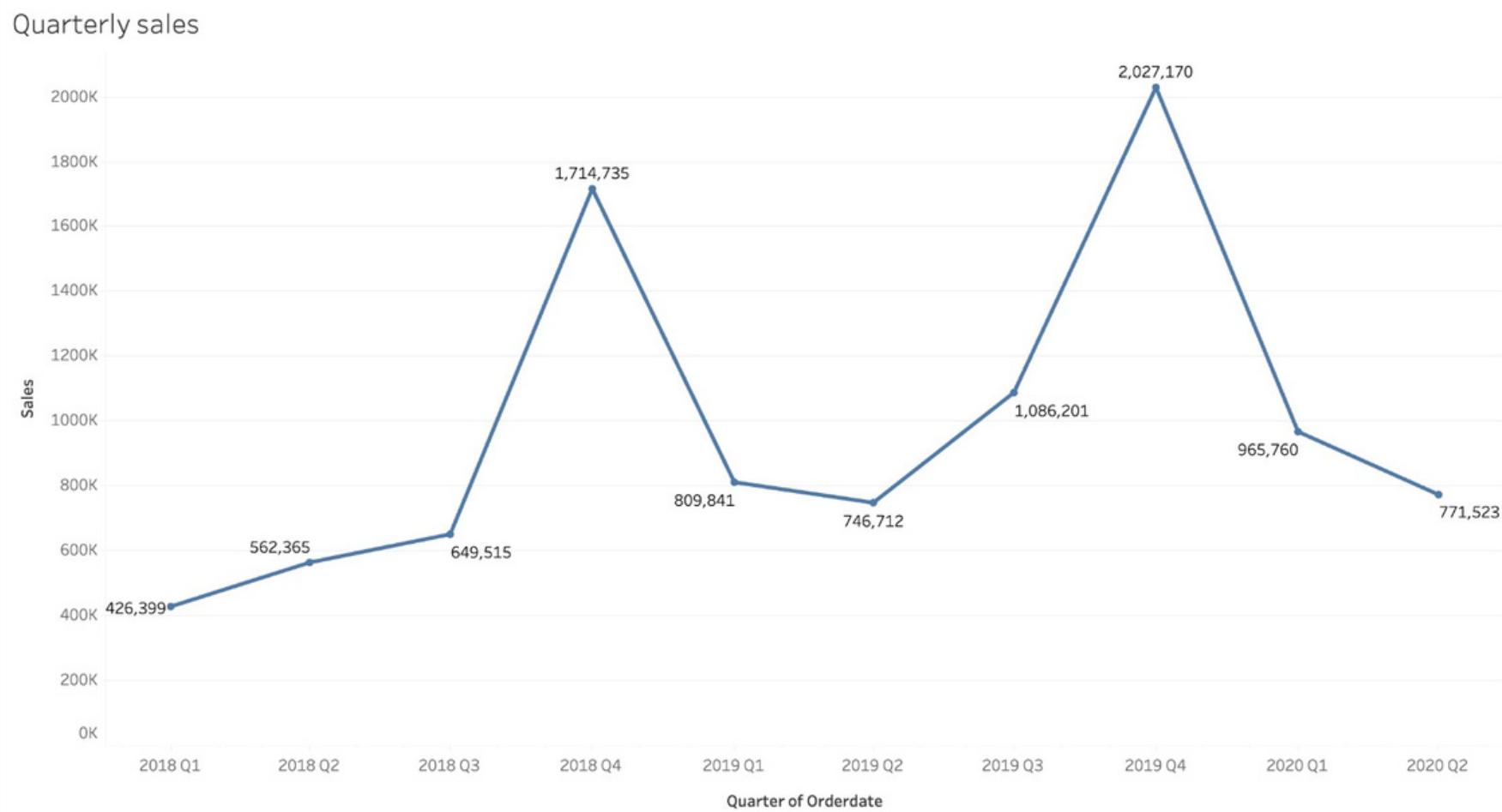
We can see an increase in sales from 2018 to 2019 and there is a sharp decline in the year of 2020 with sales dropping below the year of 2018

Monthly Sales



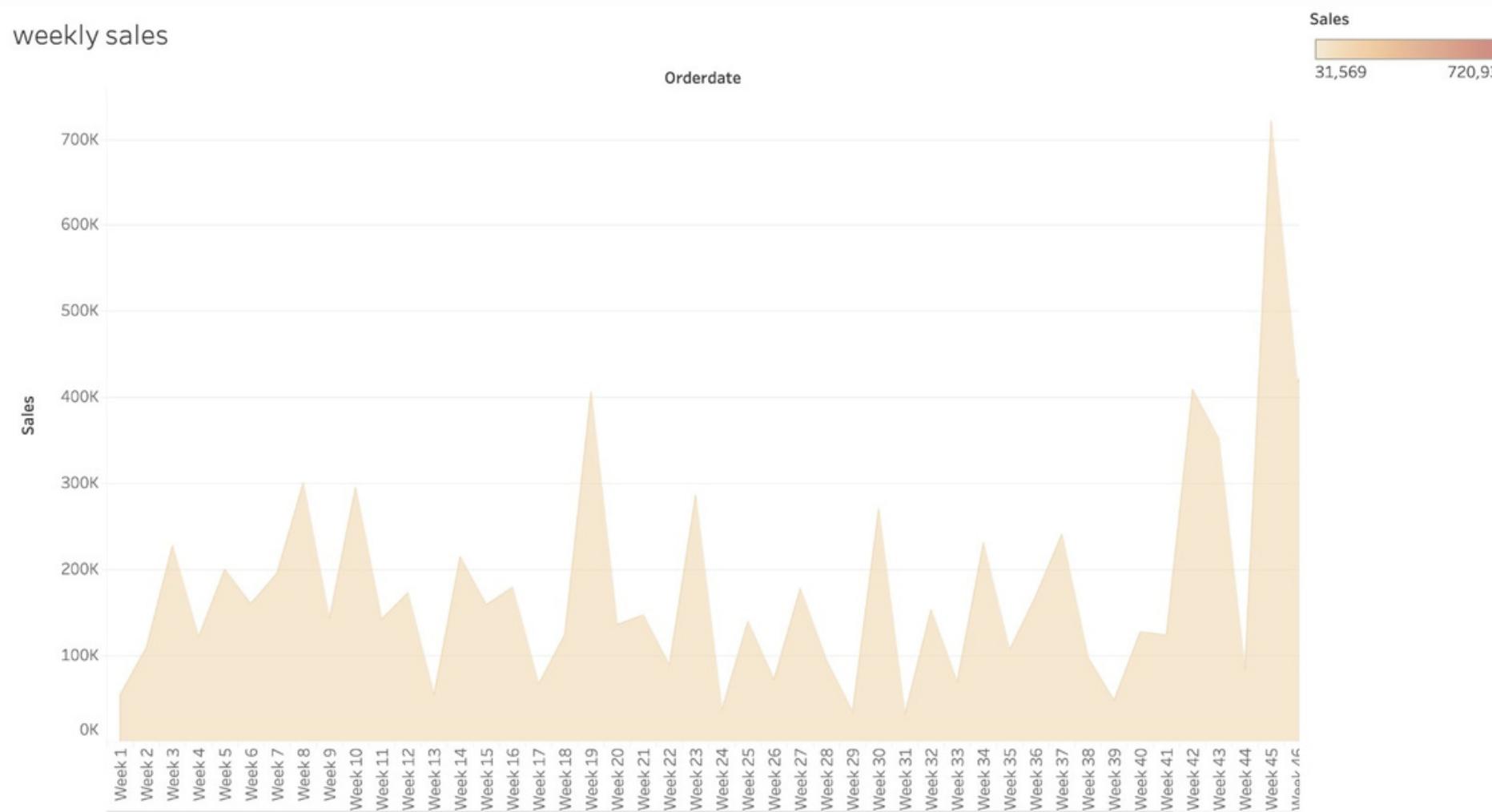
From the data, we can observe that the peak of orders were ordered during the month of October and November. This is when we must be prepared with enough demand to ship on time

Quarterly Sales



As observed here, Quarter 4 is where there are highest number of sales, followed by a sharp decline for Quarter 1 for all the 3 years of data we are seeing .

Weekly Sales



The highest of orders are observed at Week 45-46 ,followed by W43 and W19. W13,W24 are with the lowest of sales

RFM Analysis

Recency

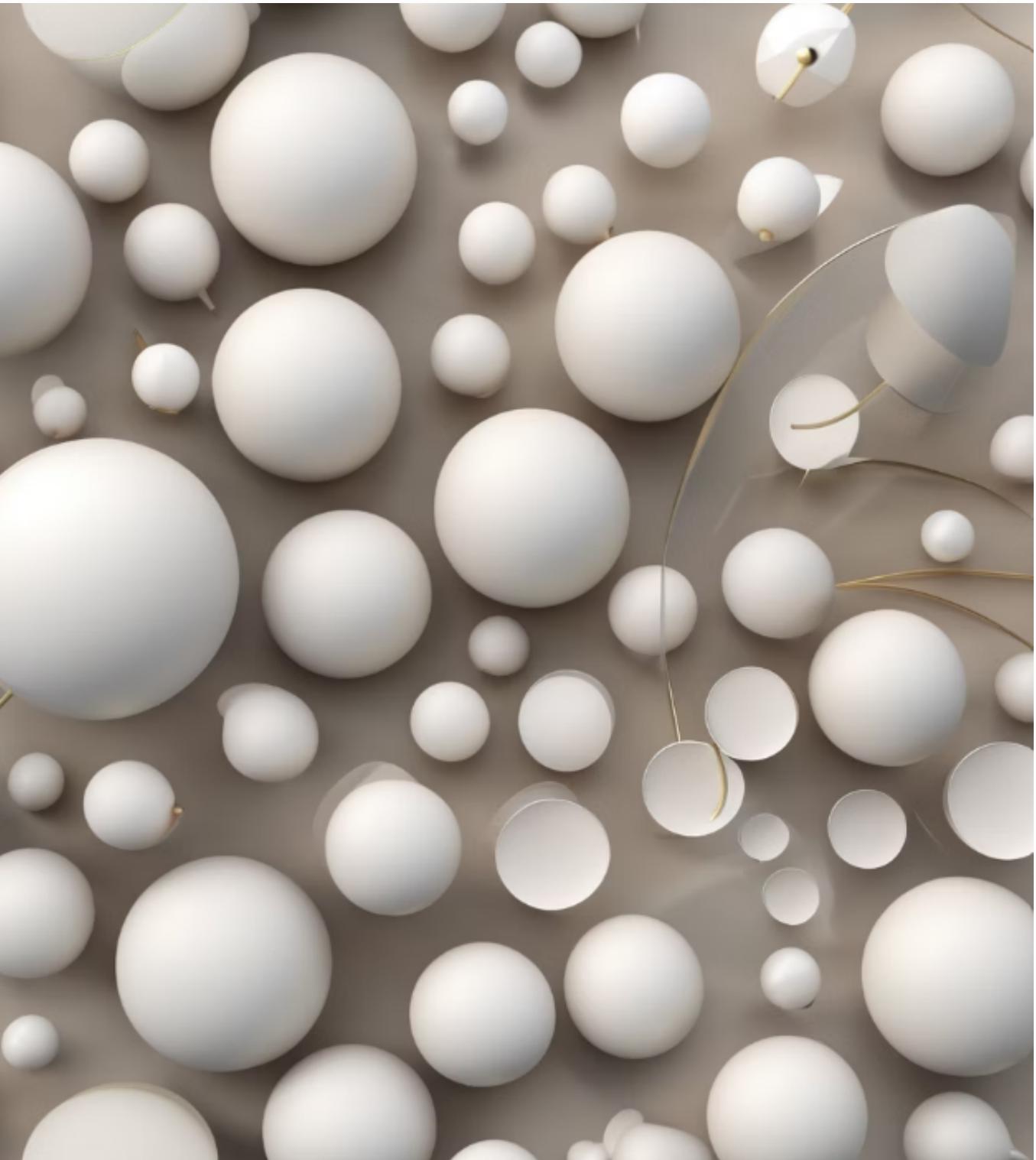
Recency refers to how recently a customer made a purchase. It helps identify the most recent customers, which can be valuable for targeted marketing campaigns.

Frequency

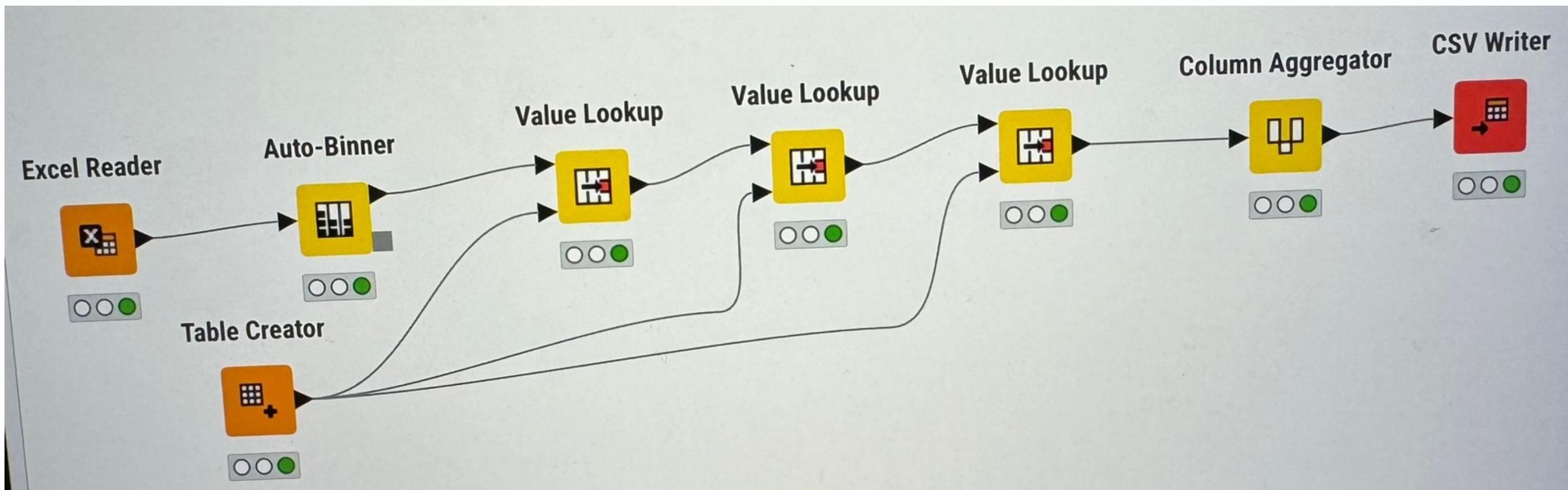
Frequency measures how often a customer makes a purchase. It helps identify loyal customers who make frequent purchases and can be targeted for retention strategies.

Monetary Value

Monetary Value refers to the amount of money a customer spends on their purchases. It helps identify high-value customers who contribute significantly to the business's revenue.



KNIME WORKFLOW



OUTPUT IMAGE

ORDERNU	QUANTITY	PRICE	EA	ORDERID	SALES	ORDERDATE	DAYS_SIN	STATUS	PRODUCT	MSRP	PRODUCT	CUSTOMERNAME	PHONE	ADDRESS	CITY	POSTALCODE	COUNTRY	CONTACT	CONTACT	DEALSIZE	QUANTITY	SALES	[B DAYS SIN]	frequency	monetary	recency	Concatenated
Order Number	Integ	Number	String	Number	Double	Number	Integ	String	String	Number	Integ	String	String	String	String	String	String	String	String	String	String	String	String	String	String	String	
10402	45	129.64	1	5,833.80			127	Shipped	Motorcycle	118	S10_2016	Auto Cana (1)	47.55.6.25, rue Lat Paris		75016	France	Perrier	Dominique	Medium	Bin 3	Bin 2	Bin 1	3	2	5		
10388	50	44.51	5	2,225.50			161	Shipped	Motorcycle	118	S10_2016	FunGiftide	5.09E+09 1785 First New Bedfo		50553	USA	Benitez	Violeta	Small	Bin 3	Bin 1	Bin 1	3	1	5		
10168	39	82.91	17	3,233.49			3255	Shipped	Planes	74	S700_4002	Technics	6.51E+09 9408 Furth Burlingame		94217	USA	Hirano	Juri	Medium	Bin 2	Bin 2	Bin 5	2	2	1		
10222	31	95.34	17	2,955.54			3120	Shipped	Ships	99	S700_3962	Collectable	7.61E+09 361 Furth (San Diego		91217	USA	Thompson	Valarie	Small	Bin 2	Bin 1	Bin 5	2	1	1		
10386	32	94.34	17	3,018.88			2704	Resolved	Planes	80	S700_3167	Euro Shop (91)	555 94 C/ Moralza Madrid		28034	Spain	Freyre	Diego	Medium	Bin 2	Bin 1	Bin 4	2	1	2		
10350	44	147.52	17	6,490.88			2536	Shipped	Trains	58	S50_1514	Euro Shop (91)	555 94 C/ Moralza Madrid		28034	Spain	Freyre	Diego	Medium	Bin 3	Bin 3	Bin 4	3	3	2		
10135	30	89.8	17	2,694			2958	Shipped	Motorcycle	102	S32_4485	Mini Gifts I	4.16E+09 5677 Stron San Rafa		97562	USA	Nelson	Valarie	Small	Bin 2	Bin 1	Bin 5	2	1	1		
10306	46	50.33	17	2,315.18			2427	Shipped	Trains	62	S32_3207	AV Stores (171)	555-1 Fauntleroy Manchester EC2 5NT	UK	Ashworth	Victoria	Small	Bin 3	Bin 1	Bin 4	3	1	2				
10310	36	43.05	17	1,549.80			2372	Shipped	Motorcycle	40	S32_2206	Toms Spe	0221-5554 Mehrheimer Kolin		50739	Germany	Pfalzheim	Henriette	Small	Bin 2	Bin 1	Bin 4	2	1	2		
10122	31	44.66	17	1,384.46			2884	Shipped	Motorcycle	40	S32_2206	Marseille	91.24.4555 12, rue des Marseille		13008	France	Lebihan	Laurence	Small	Bin 2	Bin 1	Bin 5	2	1	1		
10182	49	127.43	17	6,244.07			2430	Shipped	Classic Ca	107	S24_3432	Mini Gifts I	4.16E+09 5677 Stron San Rafa		97562	USA	Nelson	Valarie	Medium	Bin 3	Bin 3	Bin 4	3	3	2		
10165	32	145.68	17	4,661.76			2177	Shipped	Trucks and	127	S24_2300	Dragon So	+65 221 75 Bronz Sok Singapore		79903	Singapore	Natividad	Eric	Medium	Bin 2	Bin 2	Bin 4	2	2	2		
10287	34	126.48	17	4,300.32			1608	Shipped	Classic Ca	148	S18_4721	Vida Sport	0897-0345 Grenzache Gensve		1203	Switzerland	Holz	Michael	Medium	Bin 2	Bin 2	Bin 3	2	2	3		
10280	35	105.83	17	3,704.05			1519	Shipped	Vintage Ca	92	S18_4409	Amica Mo	011-498851 Via Monte Torino		10100	Italy	Accorti	Paolo	Medium	Bin 2	Bin 2	Bin 3	2	2	3		
10398	45	106.93	17	4,811.85			1253	Shipped	Vintage Ca	105	S18_3856	Reims Coll	26.47.1555 59 rue de l Reims		51100	France	Henriot	Paul	Medium	Bin 3	Bin 2	Bin 2	3	2	4		
10312	30	61.15	16	1,834.50			890	Shipped	Vintage Ca	60	S18_2248	Mini Gifts I	4.16E+09 5677 Stron San Rafa		97562	USA	Nelson	Valarie	Small	Bin 2	Bin 1	Bin 2	2	1	4		
10204	38	169.28	16	6,432.64			1153	Shipped	Classic Ca	142	S18_1984	Muscle Ma	2.13E+09 4092 Furth NYC		10022	USA	Young	Jeff	Medium	Bin 2	Bin 3	Bin 2	2	3	4		
10280	26	141.1	16	3,668.60			852	Shipped	Vintage Ca	170	S18_1749	Amica Mo	011-498851 Via Monte Torino		10100	Italy	Accorti	Paolo	Medium	Bin 2	Bin 2	Bin 2	2	2	4		
10110	37	146.84	16	5,433.08			1307	Shipped	Classic Ca	124	S18_1589	AV Stores (171)	555-1 Fauntleroy Manchester EC2 5NT	UK	Ashworth	Victoria	Medium	Bin 2	Bin 2	Bin 2	2	2	4				
10332	27	89.89	16	2,427.03			692	Shipped	Vintage Ca	53	S18_1367	AV Stores (171)	555-1 Fauntleroy Manchester EC2 5NT	UK	Ashworth	Victoria	Small	Bin 2	Bin 1	Bin 1	2	1	5				
10192	27	131.28	16	3,544.56			941	Shipped	Classic Ca	115	S12_4675	Online Die	6.04E+09 2304 Long Nashua		62005	USA	Young	Valarie	Medium	Bin 2	Bin 2	Bin 2	2	2	4		
10212	39	126.84	16	4,946.76			785	Shipped	Classic Ca	117	S12_3380	Euro Shop (91)	555 94 C/ Moralza Madrid		28034	Spain	Freyre	Diego	Medium	Bin 2	Bin 2	Bin 2	2	2	4		
10373	45	55.62	17	2,502.90			2812	Shipped	Planes	74	S700_4002	Oulu Tov	981-443651 Torikatu 3& Oulu		90110	Finland	Koskitalo	Pirkko	Small	Bin 3	Bin 1	Bin 4	3	1	2		

Inferences from RFM Analysis and identified segments

-> Who are your best customers? (give at least 5) ->
Which customers are on the verge of churning? (give at
least 5) -> Who are your lost customers? (give at least
5) -> Who are your loyal customers? (give at least 5)

CONTACTLA	CONTACTFIF	DEALSIZE	QUANTITYOF SALES [Binne	DAYS SINCE	frequency	monetary	recency	Concatenate
Citeaux	Frederique	Large	Bin 5	Bin 4	Bin 1	5	4	5
Brown	William	Large	Bin 5	Bin 3	Bin 4	5	3	2
Devon	Elizabeth	Large	Bin 4	Bin 5	Bin 1	4	5	5
Frick	Sue	Large	Bin 4	Bin 5	Bin 1	4	5	455
Citeaux	Frederique	Large	Bin 4	Bin 5	Bin 3	4	5	3

The above are the best 5 customers for the company,

they have very goof frequency rate and high monetory rate

The below are the top 5 loyal customers as their recency rate is very high along with frequency

CONTACTLA	CONTACTFIF	DEALSIZE	QUANTITYOF SALES [Binne	DAYS SINCE	frequency	monetary	recency	Concatenate
Citeaux	Frederique	Large	Bin 5	Bin 4	Bin 1	5	4	5
Devon	Elizabeth	Large	Bin 4	Bin 5	Bin 1	4	5	5
Frick	Sue	Large	Bin 4	Bin 5	Bin 1	4	5	5
Frick	Sue	Large	Bin 4	Bin 4	Bin 1	4	4	5
Devon	Elizabeth	Large	Bin 4	Bin 4	Bin 1	4	4	5

The below are the worst 5 customers who probably churned as they are inactive for a long time and their frequency and monetory rate is too bad

Donnermeyer	Michael	Small	Bin 1	Bin 1	Bin 5	1	1	1	1
Huxley	Adrian	Small	Bin 1	Bin 1	Bin 5	1	1	1	1
Bennett	Helen	Small	Bin 1	Bin 1	Bin 5	1	1	1	1
Yu	Kwai	Small	Bin 1	Bin 1	Bin 5	1	1	1	1
Freyre	Diego	Small	Bin 1	Bin 1	Bin 5	1	1	1	1

The 5 customers who are on the verge of churning are mentioned below

On the verge of churning

Devon	Elizabeth
Freyre	Diego
King	Sue
Young	Valarie
Young	Jeff

INFERENCE

- Customer retention strategies involve actively engaging and retaining customers who are at risk of leaving. This entails offering personalized incentives, loyalty programs, or exclusive promotions tailored to encourage their ongoing loyalty. By understanding individual customer preferences and behavior, businesses can develop targeted approaches to strengthen relationships and prevent churn.
- Continuous monitoring and analysis are essential components of effective customer retention efforts. Regularly tracking customer behavior, sales trends, and key performance indicators enables businesses to identify evolving patterns and proactively address potential issues before they escalate. This proactive approach allows organizations to stay agile and responsive to changing market dynamics and customer needs.
- Additionally, identifying upselling and cross-selling opportunities among existing customers is crucial for maximizing revenue and fostering loyalty. By analyzing purchase history and customer preferences, businesses can pinpoint opportunities to recommend complementary products or services that align with their interests. This not only increases the average order value but also enhances the overall customer experience and strengthens the relationship between the customer and the brand.

THANK YOU!