



# AutoMobile Part Manufacturing Company



## RFM Analysis

Submitted By:

Akhil Mohandas



# Contents

- Agenda & Executive Summary of the data:
  - Problem statement
  - About Data
- Exploratory Analysis and Inferences:
  - Univariate, Bivariate, and multivariate analysis
  - Trends in Sales
- Customer Segmentation using RFM analysis
  - What is RFM
  - Parameters used and assumptions made
  - KNIME workflow image
- Inferences from RFM Analysis and identified segments
  - Best customers
  - customers on the verge of churning
  - lost customers
  - loyal customers

# Agenda

## 1. Introduction

- Overview of the project goals and objectives.

## 2. Data Understanding and Preparation

- Summary of the dataset and key features.

- **Exploratory Data Analysis (EDA)**

- Insights into customer buying patterns and behavior.

- **Segmentation and Pattern Identification**

- Techniques used for clustering and insights derived from customer segmentation.

- **Recommendations**

- Customized marketing strategies for different customer segments.

- **Conclusion**

- Key takeaways and next steps.

# Executive Summary

- The company aims to analyze 3 years of transactional data to identify customer buying patterns and provide insights.
- The data includes 19 columns capturing transactional, product, and customer details.
- Key objectives: uncover customer segments, understand purchase frequency and volume, and recommend tailored marketing strategies.
- Deliverables: Insights report, segmentation results, and actionable recommendations.

# Problem Statement

The automobile parts manufacturing company aims to enhance its understanding of customer buying patterns and preferences. Without an in-house data science team, the company has enlisted a consultant to analyze three years of transactional data.

Key goals include:

- Identifying customer segments based on purchasing behavior.
- Uncovering trends and insights to boost sales and customer retention.
- Developing customized marketing strategies tailored to the needs of each customer segment.

Challenges:

- Large and diverse dataset requiring detailed analysis.
- No prior segmentation or analytics framework in place.
- Need for actionable insights to drive marketing and sales strategies effectively.

# Executive Summary

- This project analyzes three years of transactional data from an automobile parts manufacturing company, aiming to identify underlying customer buying patterns and recommend customized marketing strategies tailored to distinct customer segments.
- Dataset Details
  - Size: 20 columns and 2747 rows
  - Data Quality: No missing or duplicate values; a few outliers were identified and addressed.
  - Variables: The dataset encompasses transactional, product, and customer-specific attributes, such as sales, order dates, product lines, deal sizes, and customer demographics.

# Data Dictionary

- **ORDERNUMBER** : This column represents the unique identification number assigned to each order.
- **QUANTITYORDERED** : It indicates the number of items ordered in each order.
- **PRICEEACH** : This column specifies the price of each item in the order.
- **ORDERLINENUMBER** : It represents the line number of each item within an order.
- **SALES** : This column denotes the total sales amount for each order, which is calculated by multiplying the quantity ordered by the price of each item.
- **ORDERDATE** : It denotes the date on which the order was placed.
- **DAYS\_SINCE\_LASTORDER** : This column represents the number of days that have passed since the last order for each customer. It can be used to analyze customer purchasing patterns.
- **STATUS** : It indicates the status of the order, such as "Shipped," "In Process," "Cancelled," "Disputed," "On Hold," or "Resolved"
- **PRODUCTLINE** : This column specifies the product line categories to which each item belongs.
- **MSRP** : It stands for Manufacturer's Suggested Retail Price and represents the suggested selling price for each item.
- **PRODUCTCODE** : This column represents the unique code assigned to each product.
- **CUSTOMERNAME** : It denotes the name of the customer who placed the order.
- **PHONE** : This column contains the contact phone number for the customer.
- **ADDRESSLINE1** : It represents the first line of the customer's address.
- **CITY** : This column specifies the city where the customer is located.
- **POSTALCODE** : It denotes the postal code or ZIP code associated with the customer's address.
- **COUNTRY** : This column indicates the country where the customer is located.
- **CONTACTLASTNAME** : It represents the last name of the contact person associated with the customer.
- **CONTACTFIRSTNAME** : This column denotes the first name of the contact person associated with the customer.
- **DEALSIZE** : It indicates the size of the deal or order, which are the categories "Small," "Medium," or "Large."



# Statistical Summary

	count	unique	top	freq	mean	min	25%	50%	75%	max	std
ORDERNUMBER	2747.0	NaN	NaN	NaN	10259.761558	10100.0	10181.0	10264.0	10334.5	10425.0	91.877521
QUANTITYORDERED	2747.0	NaN	NaN	NaN	35.103021	6.0	27.0	35.0	43.0	97.0	9.762135
PRICEEACH	2747.0	NaN	NaN	NaN	101.098951	26.88	68.745	95.55	127.1	252.87	42.042548
ORDERLINENUMBER	2747.0	NaN	NaN	NaN	6.491081	1.0	3.0	6.0	9.0	18.0	4.230544
SALES	2747.0	NaN	NaN	NaN	3553.047583	482.13	2204.35	3184.8	4503.095	14082.8	1838.953901
ORDERDATE	2747	NaN	NaN	NaN	2019-05-13 21:56:17.211503360	2018-01-06 00:00:00	2018-11-08 00:00:00	2019-06-24 00:00:00	2019-11-17 00:00:00	2020-05-31 00:00:00	NaN
DAYS_SINCE_LASTORDER	2747.0	NaN	NaN	NaN	1757.085912	42.0	1077.0	1761.0	2436.5	3562.0	819.280576
STATUS	2747	6	Shipped	2541	NaN	NaN	NaN	NaN	NaN	NaN	NaN
PRODUCTLINE	2747	7	Classic Cars	949	NaN	NaN	NaN	NaN	NaN	NaN	NaN
MSRP	2747.0	NaN	NaN	NaN	100.691664	33.0	68.0	99.0	124.0	214.0	40.114802
PRODUCTCODE	2747	109	S18_3232	51	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CUSTOMERNAME	2747	89	Euro Shopping Channel	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
PHONE	2747	88	(91) 555 94 44	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
ADDRESSLINE1	2747	89	C/ Moralzarzal, 86	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CITY	2747	71	Madrid	304	NaN	NaN	NaN	NaN	NaN	NaN	NaN
POSTALCODE	2747	73	28034	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
COUNTRY	2747	19	USA	928	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CONTACTLASTNAME	2747	76	Freyre	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CONTACTFIRSTNAME	2747	72	Diego	259	NaN	NaN	NaN	NaN	NaN	NaN	NaN
DEALSIZE	2747	3	Medium	1349	NaN	NaN	NaN	NaN	NaN	NaN	NaN



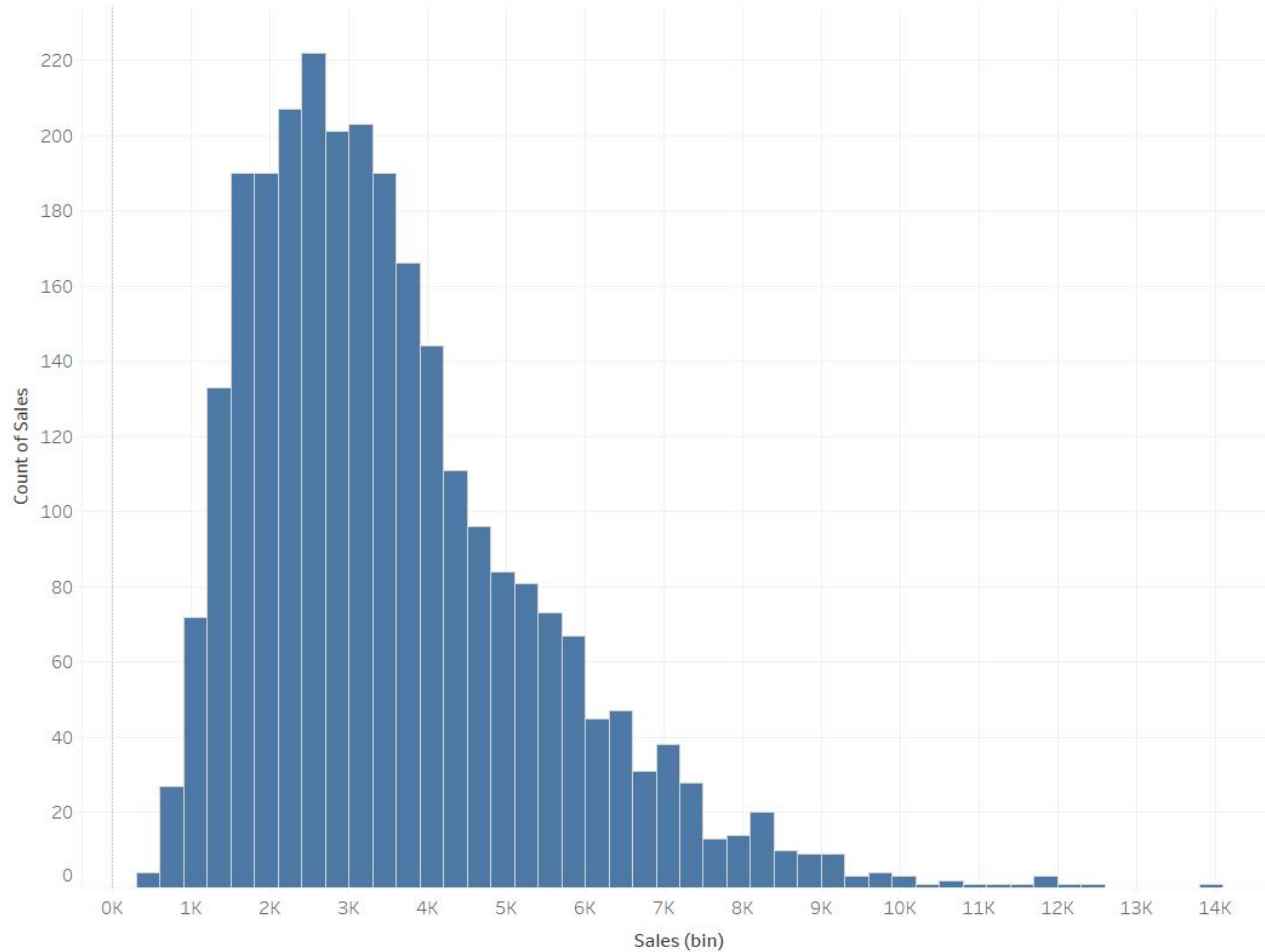
# Inference

- The ORDERNUMBER column indicates 2747 transactions with a consistent range (min: 10100, max: 10425), suggesting a steady flow of transactions over the three years.
- The average sales per order is 3553.05 dollars, with substantial variability (std dev: \$1838.95).
- Some customers appear inactive, while others might show repeat purchasing behavior within shorter intervals.
- "Classic Cars" is the most frequently purchased category
- The most frequent customer (CUSTOMERNAME = "Euro Shopping Channel") has 259 orders, suggesting a significant B2B relationship.

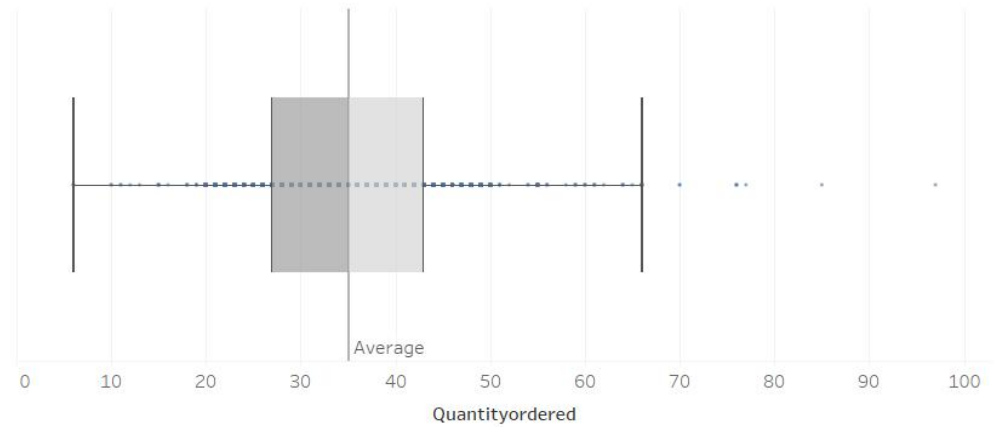
# Exploratory Data Analysis

# Univariate Analysis

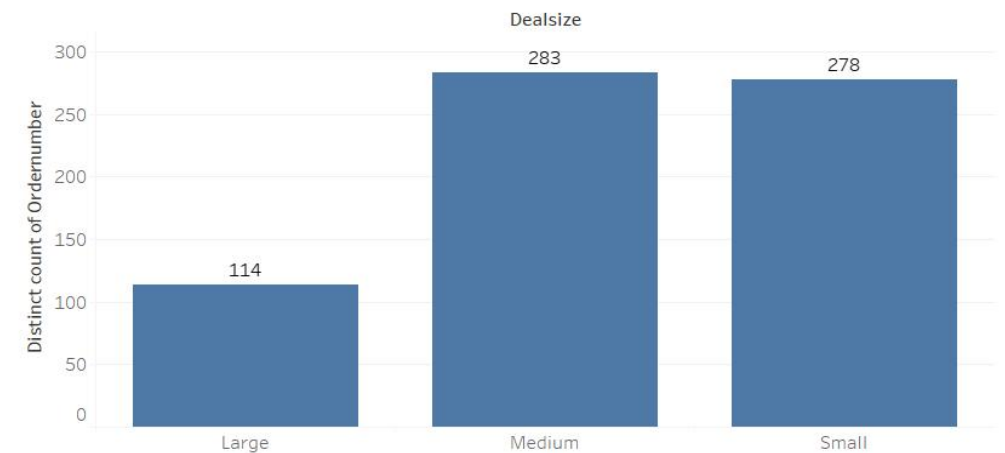
Distribution of SALES



Quantity Ordered



Dealsize Proportions

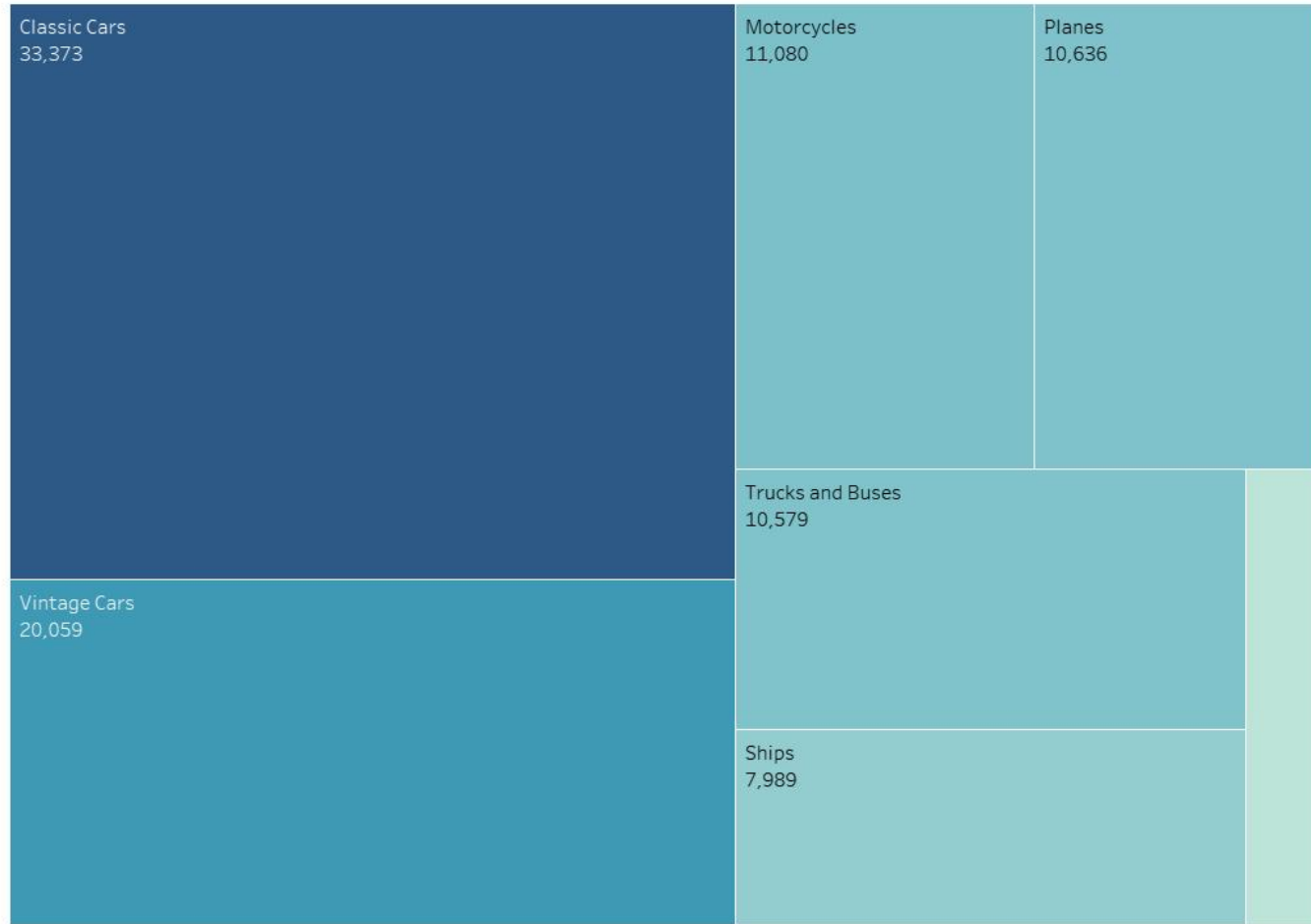


# Inferences

- We can see that sales data is skewed toward left.
- There are many outliers for Quantity ordered.
- Most data exist between 1.5k - 3.5k.
- Most sales are for medium and small while only 16% of the sales is Large.

# Univariate Analysis

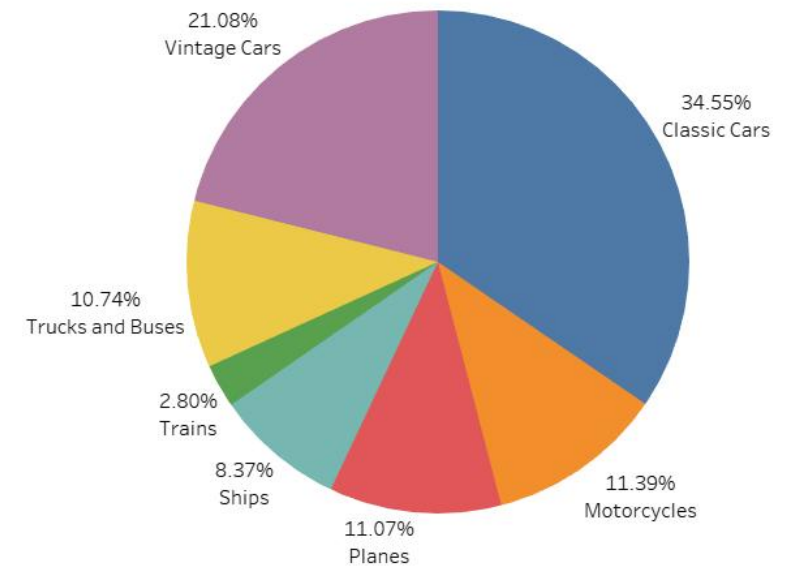
Product Line Quantity Ordered



Quantityordered



Productline Pie Chart

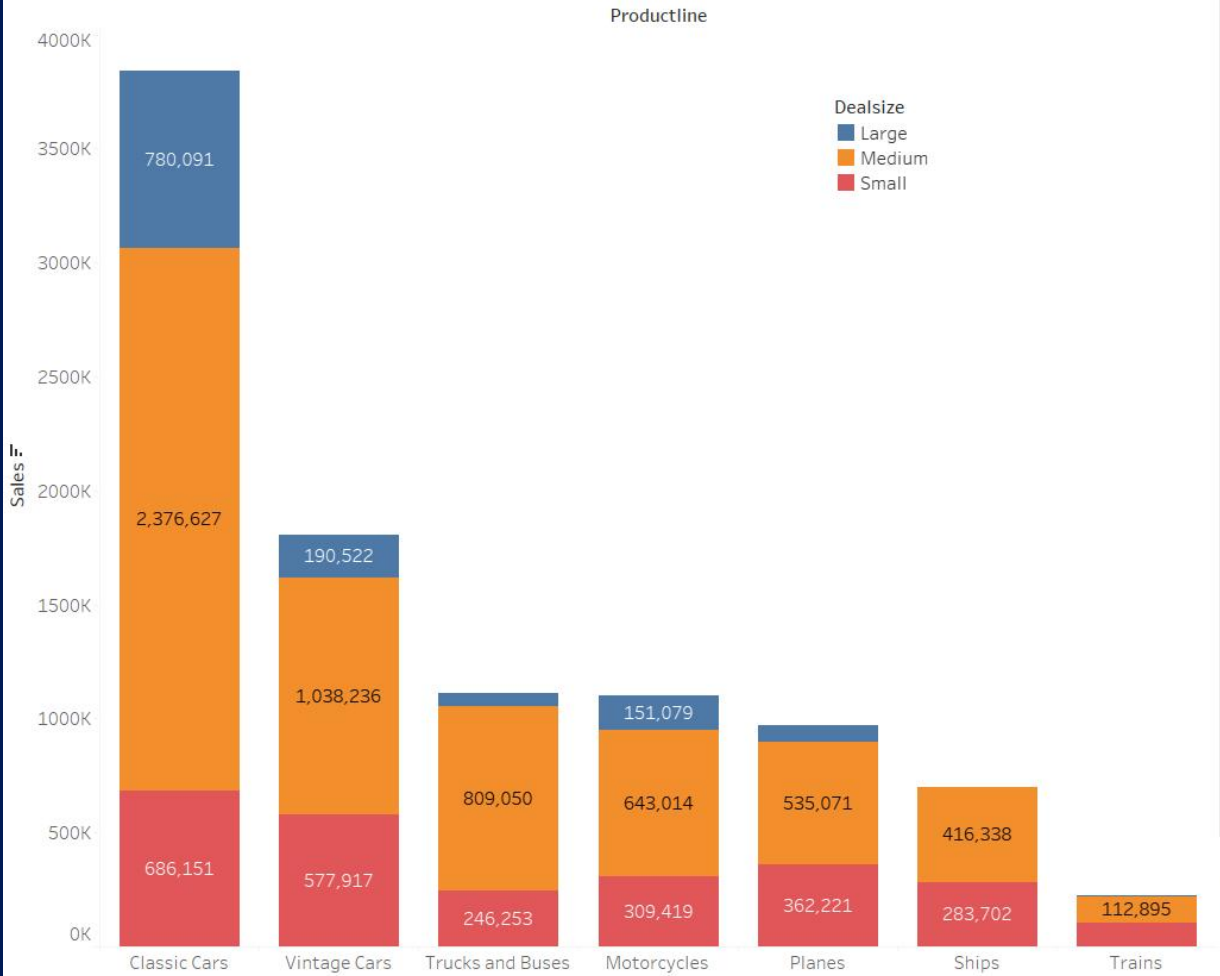


# Inferences

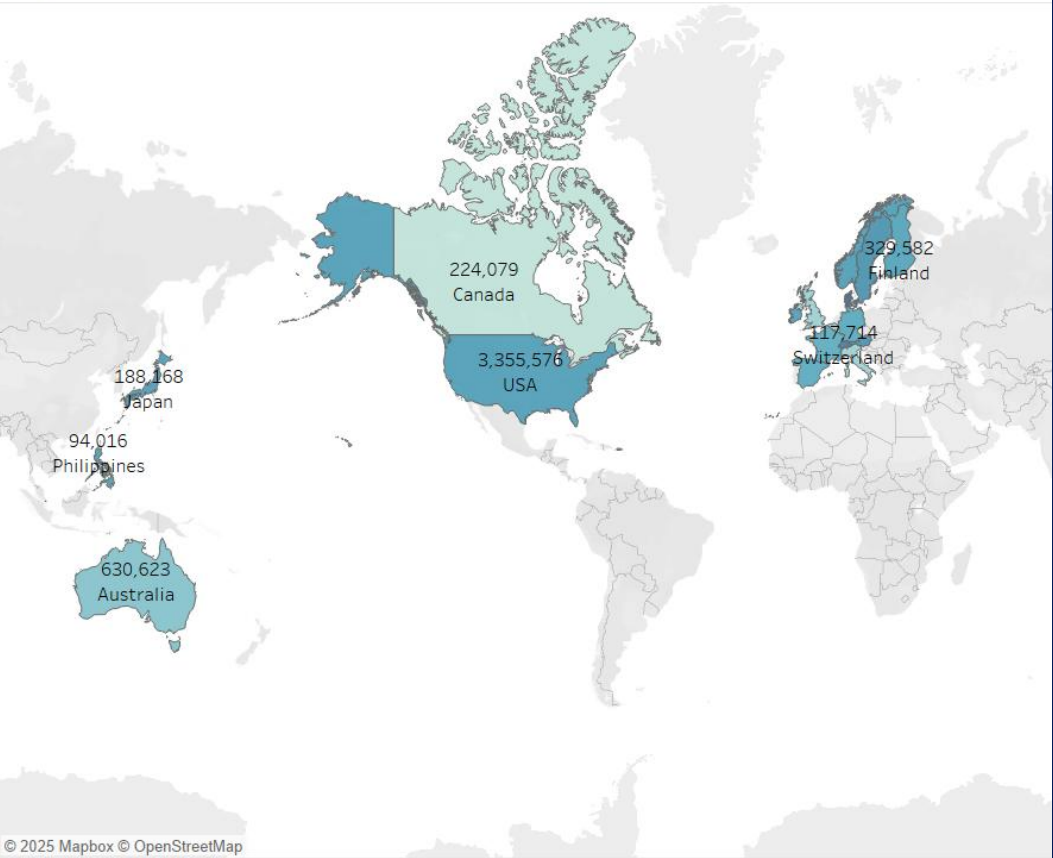
- Automobile parts for Classic cars are sold more with vintage cars coming in second.
- Train parts are sold the least.
- 50% of the product line is composed of Classic cars and Vintage cars.
- Trains and Ships comprise only approx 10% of the product line.

# Bivariate Analysis

Sales by Product Line



Avg. Sales by Country



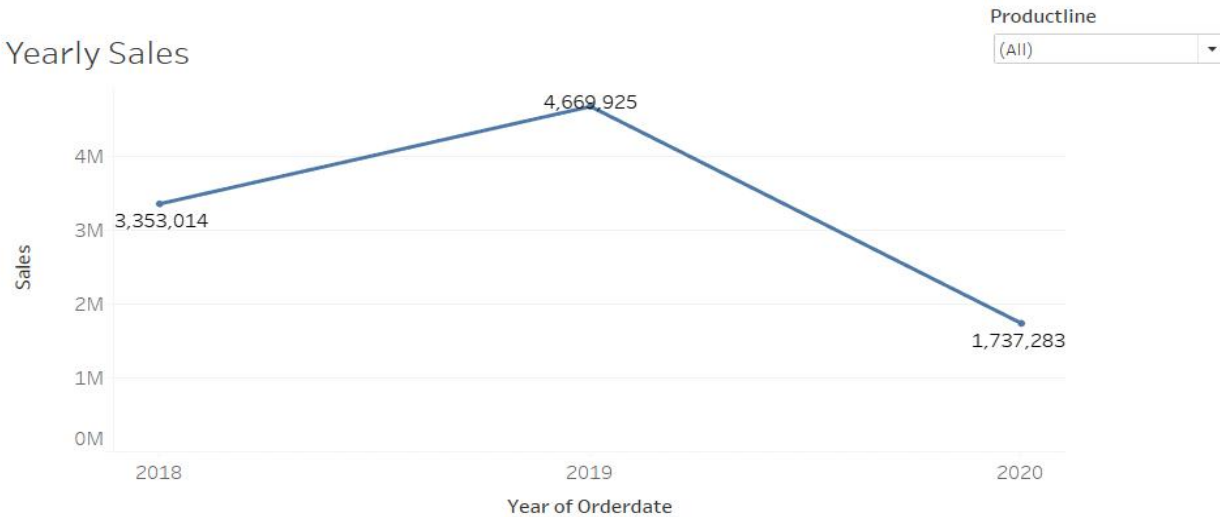


# Inferences

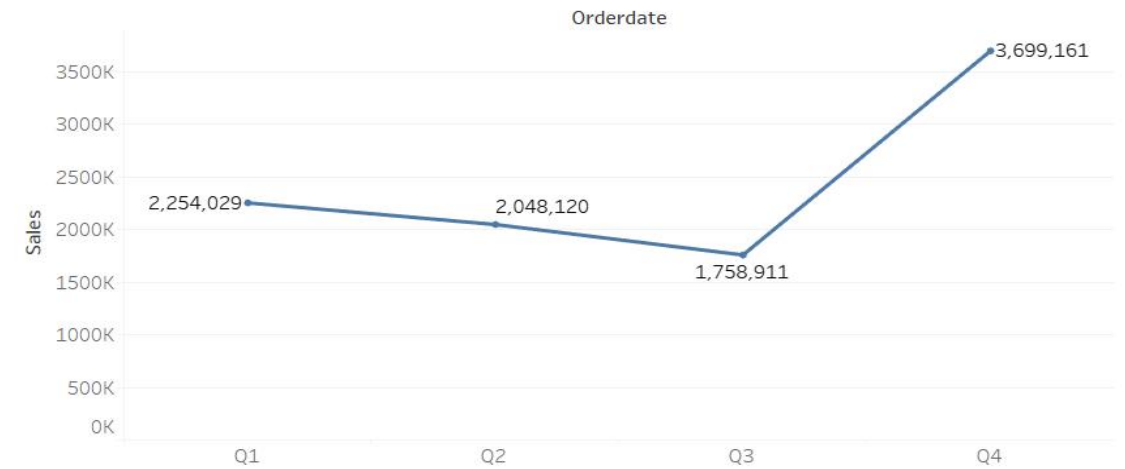
- Classic cars parts pulls maximum sales with medium deal size.
- Trains and Ships have low to none sales for large deal size.
- Maximum sales are from USA and Philippines the least.
- Medium deal size are higher for every kind of product line.

# Sales Trend

## Yearly Sales



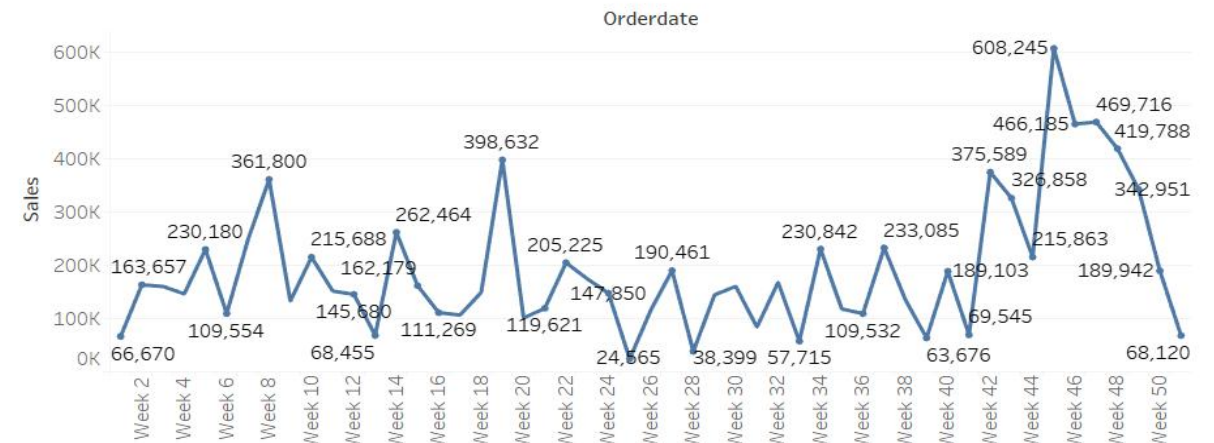
## Quarterly Sales



## Monthly Sales



## Weekly Sales



# Inferences

- Drastic drop in sales for year 2020, might be caused due to covid.
- November shows the maximum sales.
- Maximum sales are for Quarter 4.
- Quarter 3 has the lowest sales.
- week 45 shows maximum sales.
- There is a sudden rise of sales in May month while comparing to the previous months.
- Sales are consistent in the first 4 months of the year.

# Recommendations

- Introduce promotional campaigns targeting small and medium deal sizes to encourage upselling to larger deal sizes.
- Provide bundled offers or discounts for large deals to attract more customers in this category.
- Review and investigate outlier transactions to determine if they reflect genuine demand or potential data entry errors.
- If valid, consider offering bulk purchase incentives for these customers.
- Invest in marketing and inventory for Classic and Vintage car parts to capitalize on demand.
- Analyze the low performance of train and ship parts to identify potential demand gaps or market opportunities.
- Tailor marketing campaigns and loyalty programs around medium deals, offering value-added services or discounts.
- Expand market presence in high-performing regions like the USA with localized campaigns and distribution centers.
- Plan marketing campaigns and inventory stock-ups in anticipation of Q4 and November demand.
- Introduce seasonal discounts or promotions during Q3 to boost sales during this low-performing period.
- Investigate the reason behind this spike (e.g., product launches or seasonal demand) and replicate successful strategies.

# Customer Segmentation using RFM analysis

# What is RFM

- RFM (Recency, Frequency, Monetary) analysis is a customer segmentation technique used in marketing to identify and target specific customer groups based on their purchasing behavior. The three parameters are:
  - Recency (R): How recently a customer made a purchase. Customers who purchased recently are more likely to buy again.
  - Frequency (F): How often a customer makes purchases. Frequent buyers are more engaged and valuable.
  - Monetary (M): How much money a customer spends. Higher spending indicates higher value.

# Parameters and Assumptions

- Parameters Used:

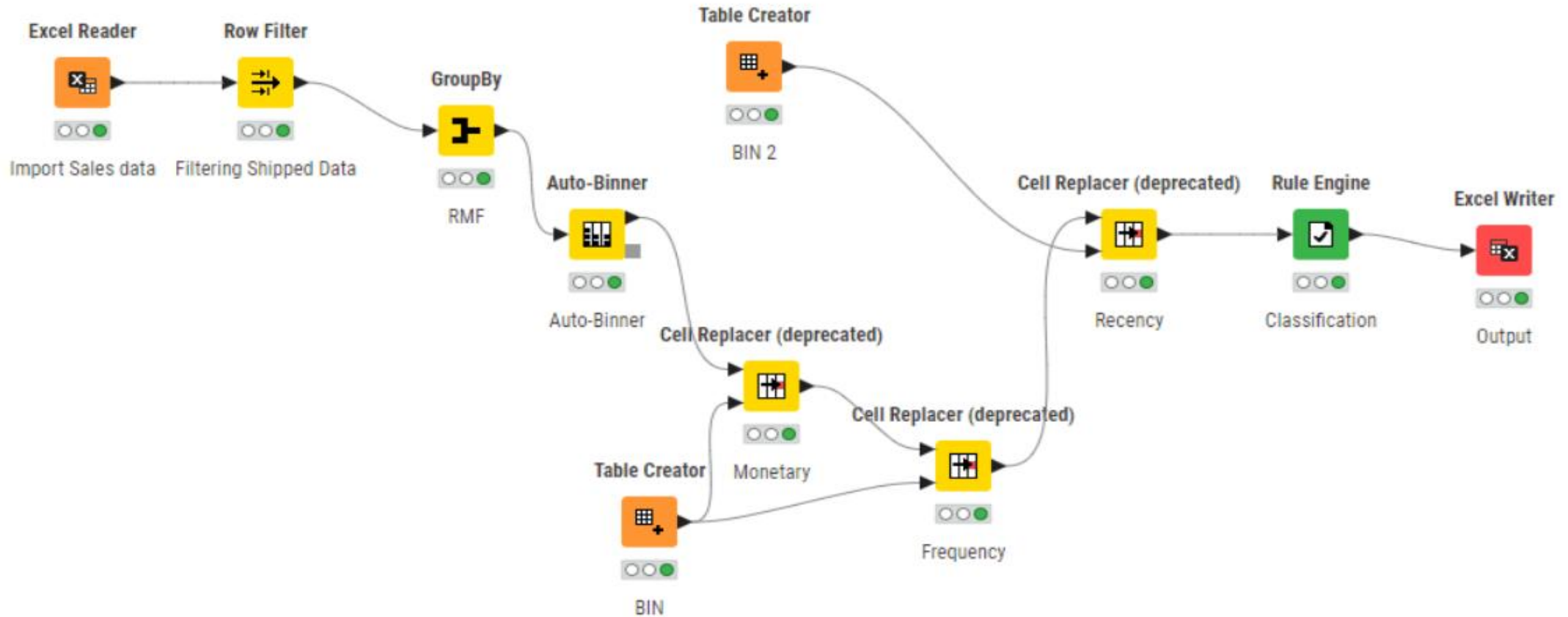
- DAYS\_SINCE\_LASTORDER → Recency: Time since the last purchase.
- Number of Orders → Frequency: Total count of orders placed by a customer.
- SALES → Monetary: Total spending by a customer(Sum of Sales).

- Assumptions:

- Only completed orders (STATUS = "Shipped") are considered.
- Recency is calculated based on DAYS\_SINCE\_LASTORDER.
- Customers are segmented into four groups(1-4, with 4 being highest and 1 lowest) for each R, F, and M, and the overall RFM score combines these segments.



# KNIME WORKFLOW



# Inferences from RFM Analysis and identified segments

# Top 5 Customers

CUSTOMERNAME	Monetary	Frequency	Recency	prediction
Mini Gifts Distributors Ltd.	4	4	4	Best Customers
Australian Collectors, Co.	4	4	4	Best Customers
Anna's Decorations, Ltd	4	4	4	Best Customers
Salzburg Collectables	4	4	4	Best Customers
La Rochelle Gifts	4	4	4	Best Customers

# Top Loyal Customers

CUSTOMERNAME	Monetary	Frequency	Recency	prediction
Baane Mini Imports	4	3	4	Loyal Customers
Corrida Auto Replicas, Ltd	4	3	3	Loyal Customers
Diecast Classics Inc.	4	3	4	Loyal Customers
Euro Shopping Channel	4	4	3	Loyal Customers
Reims Collectables	4	4	3	Loyal Customers
Tokyo Collectables, Ltd	4	3	4	Loyal Customers
Amica Models & Co.	3	3	3	Loyal Customers
Auto Canal Petit	3	3	4	Loyal Customers
FunGiftIdeas.com	3	3	4	Loyal Customers
Gift Depot Inc.	3	2	4	Loyal Customers
Heintze Collectables	3	3	3	Loyal Customers
La Corne D'abondance, Co.	3	2	3	Loyal Customers
Marta's Replicas Co.	3	3	3	Loyal Customers
Mini Classics	3	3	3	Loyal Customers
Oulu Toy Supplies, Inc.	3	3	4	Loyal Customers
Scandinavian Gift Ideas	3	2	3	Loyal Customers
Souvenirs And Things Co.	3	4	4	Loyal Customers
Suominen Souvenirs	3	3	3	Loyal Customers
The Sharp Gifts Warehouse	3	3	4	Loyal Customers
Toms Spezialitten, Ltd	3	3	3	Loyal Customers
Toys of Finland, Co.	3	3	4	Loyal Customers

# Verge on Churning

CUSTOMERNAME	Monetary	Frequency	Recency	prediction
Atelier graphique	1	1	3	At-Risk
Australian Gift Network, Co	1	1	4	At-Risk
Auto Assoc. & Cie.	1	1	3	At-Risk
Boards & Toys Co.	1	1	3	At-Risk
CAF Imports	1	1	2	At-Risk

# Lost Customers

CUSTOMERNAME	Monetary	Frequency	Recency	prediction
Australian Collectables, Ltd	1	1	1	Lost Customers
Auto-Moto Classics Inc.	1	1	1	Lost Customers
Bavarian Collectables Imports, Co.	1	1	1	Lost Customers
Clover Collections, Co.	1	1	1	Lost Customers
Double Decker Gift Stores, Ltd	1	1	1	Lost Customers
Iberia Gift Imports, Corp.	1	1	1	Lost Customers
Mini Auto Werke	1	1	1	Lost Customers
Royale Belge	1	1	1	Lost Customers
Signal Collectibles Ltd.	1	1	1	Lost Customers
Toys4GrownUps.com	1	1	1	Lost Customers

# Recommendations

## 1. Best Customers:

- Characteristics: Recent buyers with high purchase frequency and high monetary value.
- Recommendations:
  - Personalized Recognition: Send thank-you notes, exclusive loyalty rewards, or recognition in newsletters.
  - Exclusive Offers: Provide early access to sales or premium products/services.
  - VIP Treatment: Create a premium loyalty tier with enhanced benefits (e.g., free shipping, dedicated support).
  - Incentives: Offer personalized discounts on their favorite products or services.

## 2. Loyal Customers:

- Characteristics: Regular buyers with moderate recency and monetary value.
- Recommendations:
  - Periodic Discounts: Offer targeted promotions or seasonal discounts.
  - Upselling Opportunities: Suggest complementary products or services to increase spending.
  - Engagement Programs: Introduce loyalty programs with achievable milestones to reward their continued support.
  - Surprise Offers: Send unexpected offers to create delight and move them into the best customers category.



# Recommendations

## 3. At-Risk:

- Characteristics: Reduced purchase frequency or monetary value, longer recency gap.
- Recommendations:
  - Customer Feedback Surveys: Identify their dissatisfaction or concerns.
  - Personalized Communication: Re-engage with tailored messages based on their previous purchases or preferences.
  - Incentives to Return: Provide attractive discounts or cashback offers for their next purchase.
  - Proactive Support: Reach out with dedicated support to resolve any issues they might have experienced.

## 4. Lost Customers:

- Characteristics: Low frequency and monetary value, with very long recency gaps.
- Recommendations:
  - Behavior Analysis: Study their purchase history to understand the reasons for disengagement.
  - Exit Offers: Provide a "We Miss You" campaign with special discounts or freebies.
  - Re-engagement Strategy: Send targeted ads or emails showcasing new products/services.
  - Churn Prevention Insights: Use the insights gained from lost customers to improve policies and practices, minimizing future attrition.

Thank You