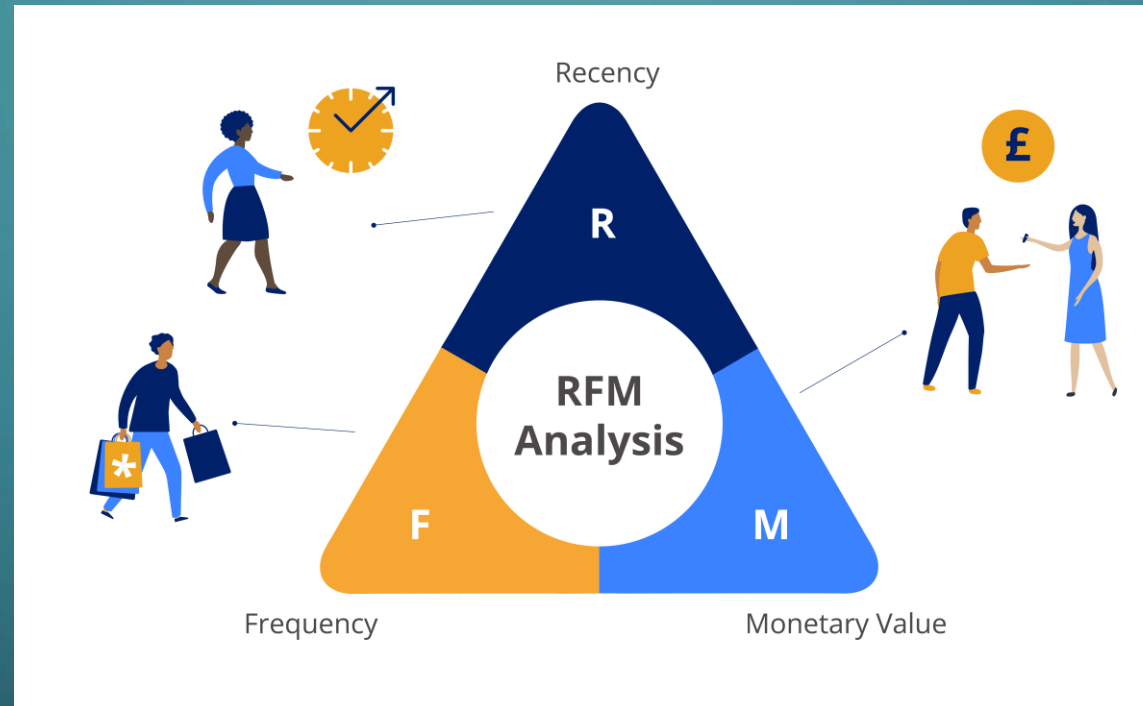


MARKET RESEARCH ANALYSIS PROJECT




RFM ANALYSIS

- BY
SUKANYA

AGENDA – CONTENTS

- Data Dictionary
- Executive Summary
- Problem Statement
- About the data
- Assumptions about the data
- Exploratory Data Analysis
 - Univariate Analysis
 - Inferences from Univariate Analysis
 - Bivariate Analysis
 - Inferences from Bivariate Analysis
 - Multivariate Analysis
 - Inferences from Multivariate Analysis
 - Yearly, Quarterly, Monthly and Weekly Trends in Sales
 - Inferences from Yearly, Quarterly, Monthly and Weekly Analysis

- 
- The slide features a dark teal background with white decorative circuit-like lines in the corners. These lines consist of vertical and horizontal segments connected by small circles, resembling a stylized electronic circuit.
- Customer Segmentation using RFM Analysis
 - What is RFM?
 - What are the parameters used and assumptions made?
 - Which tool is used for RFM? - KNIME Workflow
 - Calculations used
 - RFM Output Table
 - Inferences from RFM Analysis
 - Best Customers
 - Customers on the verge of churning
 - Lost Customers
 - Loyal Customers

DATA DICTIONARY

Column Name	Description
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."

EXECUTIVE SUMMARY OF THE DATA

- 9 customers are the best customers as they have the highest recency, frequency and monetary.
- 23 customers are active customers.
- 44 customers are at risk and can be on the verge of churning.
- 12 customers with the lowest recency, frequency and monetary value do not add anything to the sales and can be removed.
- Some of the best customers are Euro Shopping Channel, Mini Gift Distributors Ltd. etc.

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.

ABOUT THE DATA

- **SHAPE OF THE DATA:** 2747 rows and 20 columns
- **INFO ABOUT THE DATA:** 8 Numerical fields and 12 Categorical fields
- NO MISSING VALUES
- **SUMMARY STATS:**

	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

For Numerical Fields

	count	unique	top	freq
STATUS	2747	6	Shipped	2541
PRODUCTLINE	2747	7	Classic Cars	949
PRODUCTCODE	2747	109	S18_3232	51
CUSTOMERNAME	2747	89	Euro Shopping Channel	259
PHONE	2747	88	(91) 555 94 44	259
ADDRESSLINE1	2747	89	C/ Moralezarzal, 86	259
CITY	2747	71	Madrid	304
POSTALCODE	2747	73	28034	259
COUNTRY	2747	19	USA	928
CONTACTLASTNAME	2747	76	Freyre	259
CONTACTFIRSTNAME	2747	72	Diego	259
DEALSIZE	2747	3	Medium	1349

For Categorical Fields

ASSUMPTIONS ABOUT THE DATA

The mean and median of the data is almost the same showing the data is not skewed.

Quantity ordered goes as low as 35 and as high as 97.

The average sales is \$3553.

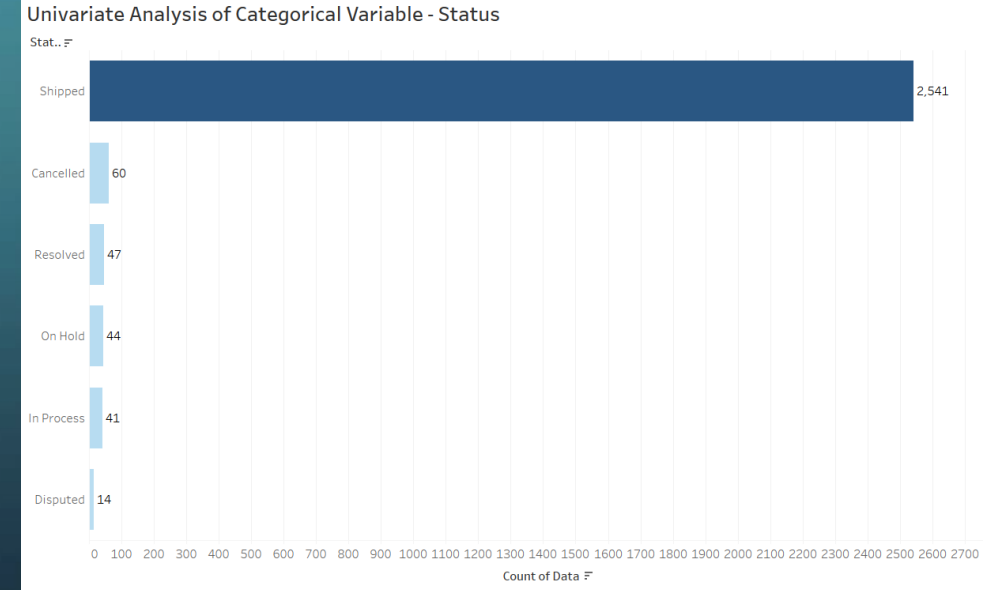
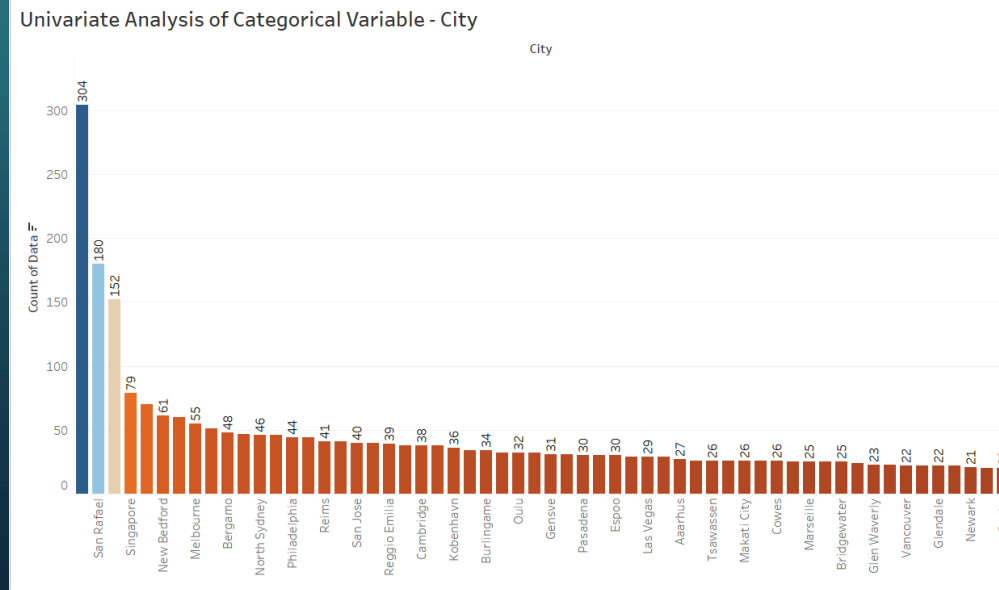
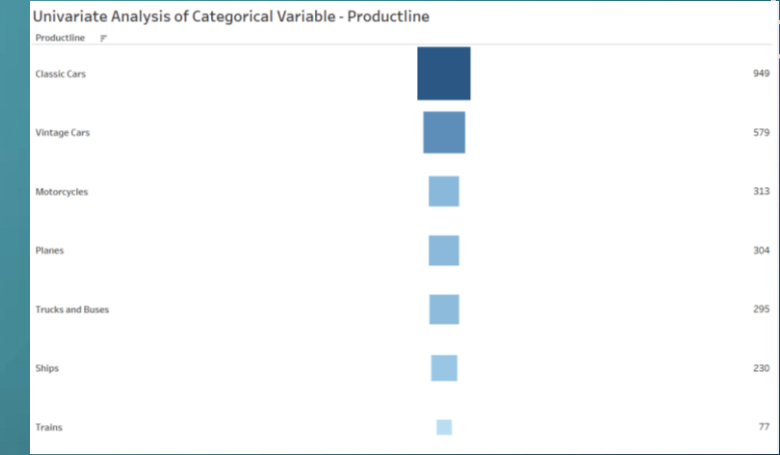
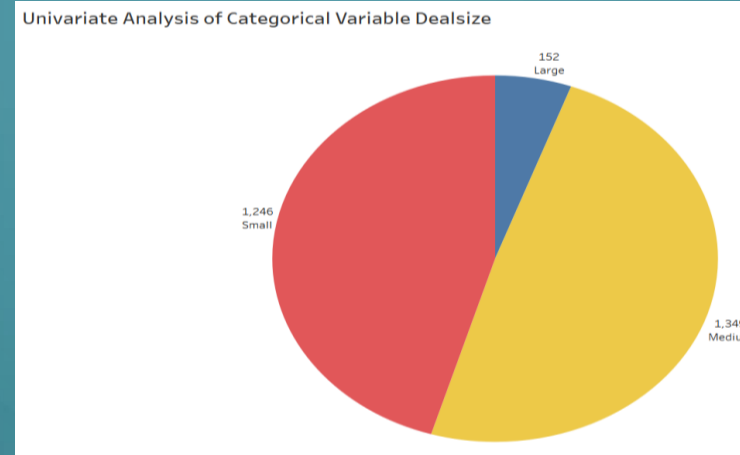
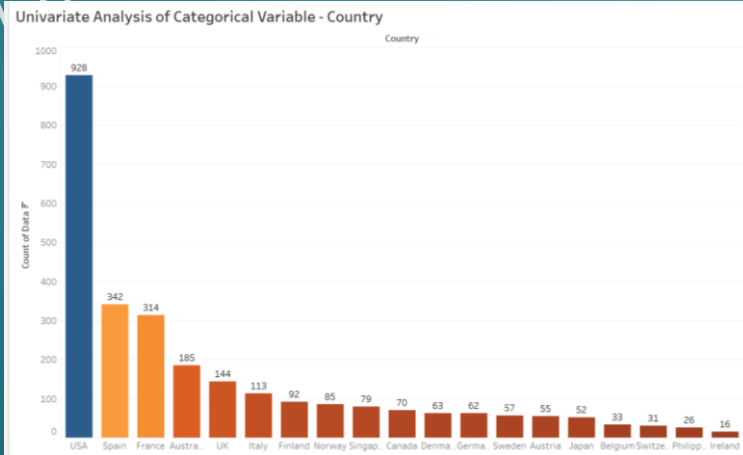
The top city in sales is Madrid while the top country is USA.

Medium deal size shows maximum sales.

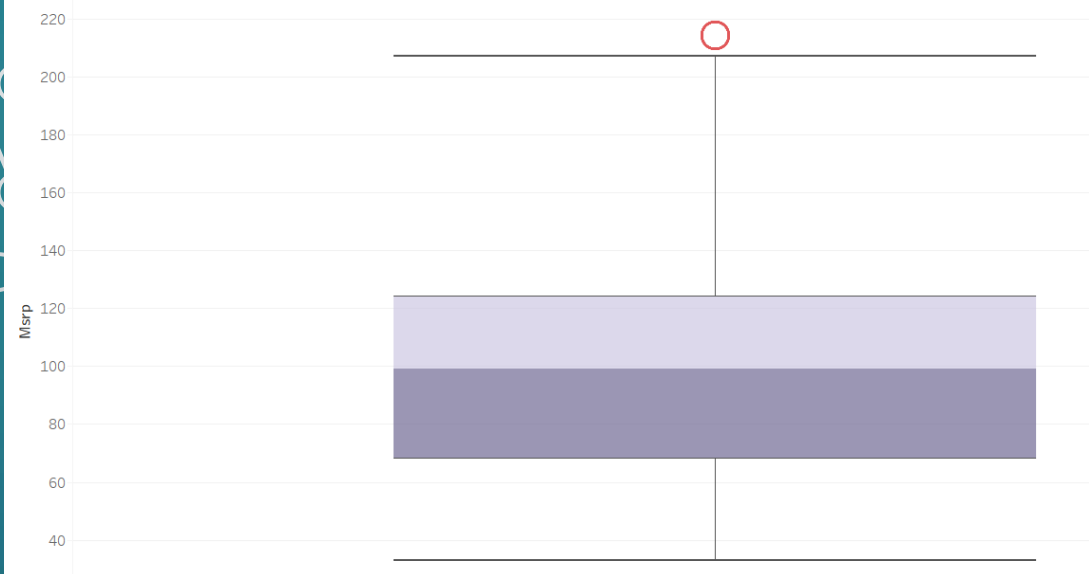
Classic cars have the highest sales among all the products.

EXPLORATORY DATA ANALYSIS

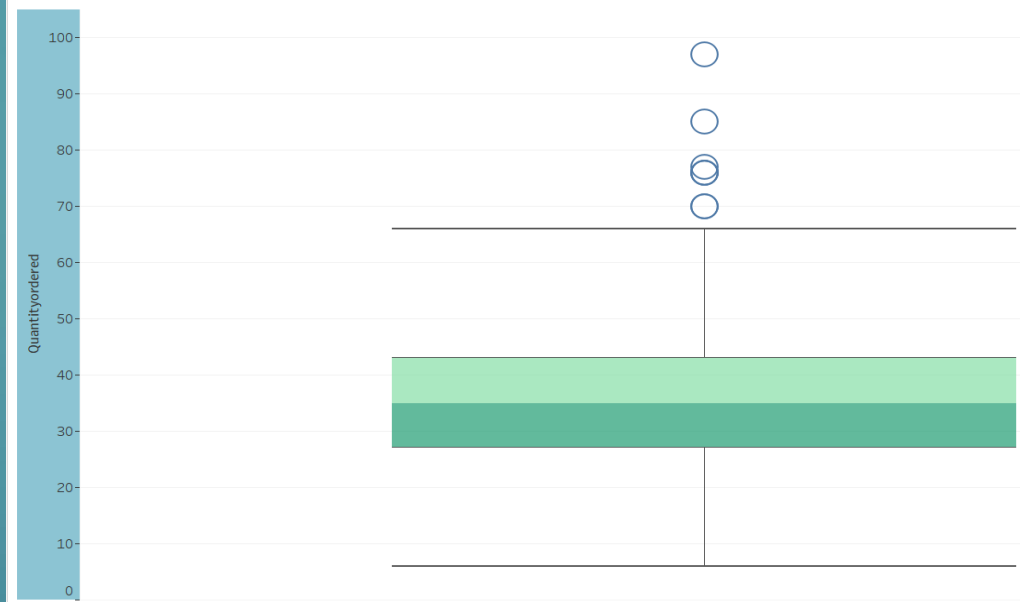
UNIVARIATE ANALYSIS



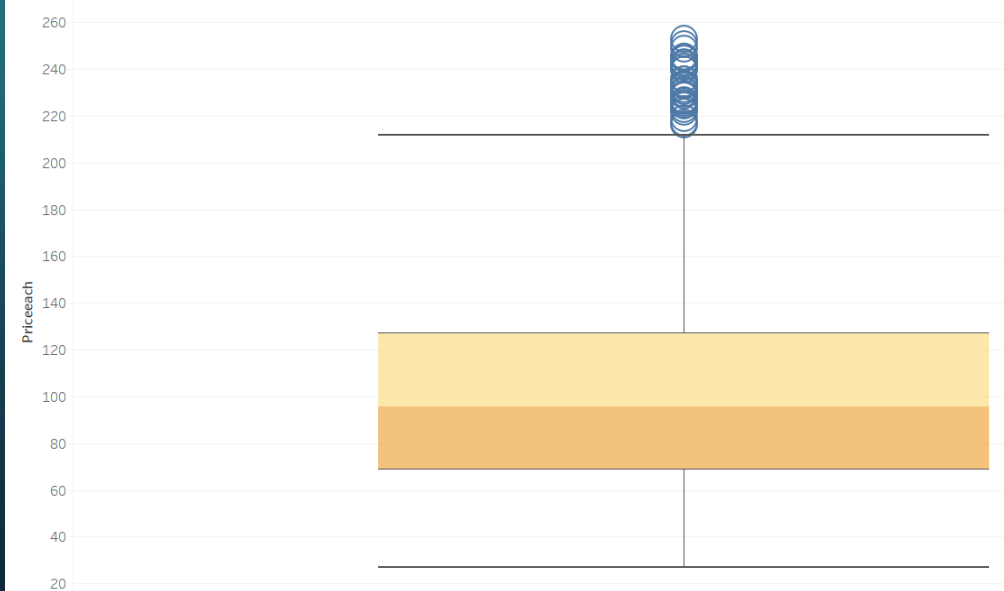
Univariate Analysis of MSRP



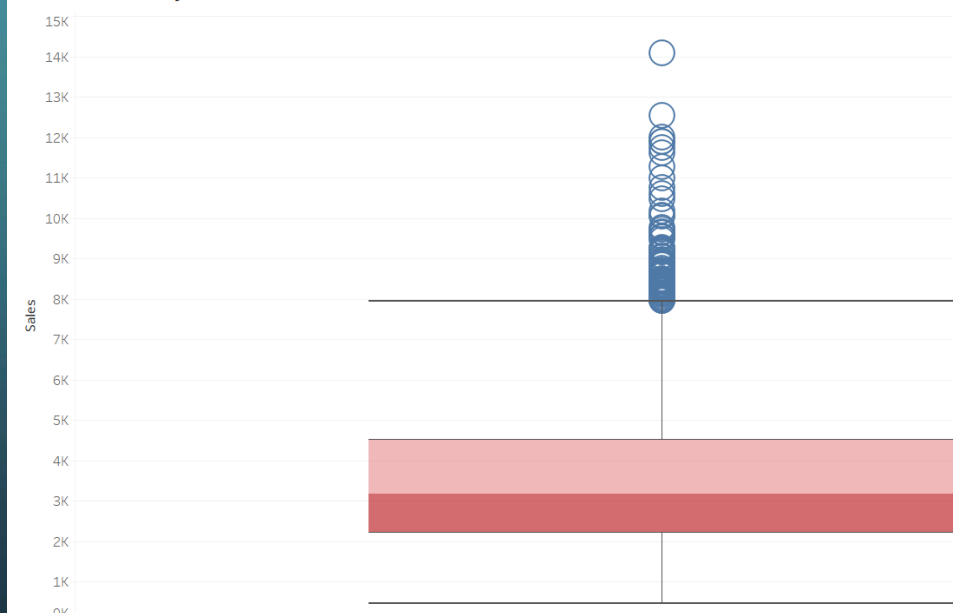
Univariate Analysis of Numerical Variable - Quantityordered



Univariate Analysis of Numerical Variable - Priceeach



Univariate Analysis of Numerical variable - Sales

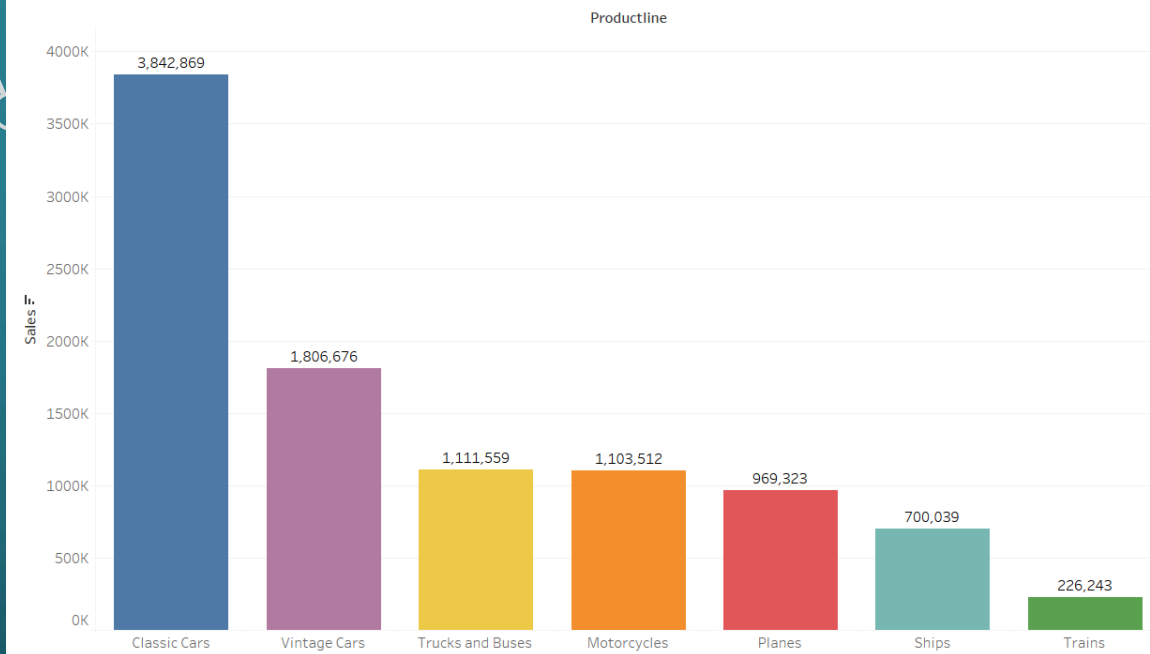


INFERENCES FROM UNIVARIATE ANALYSIS:

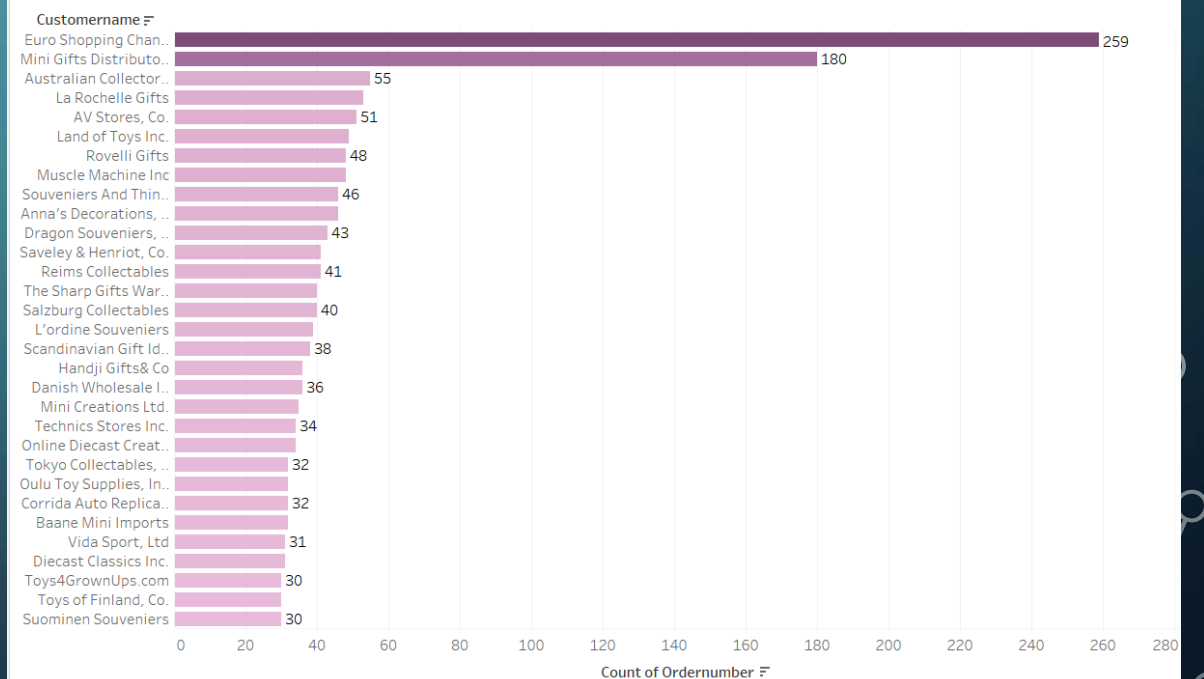
- USA has the highest count of sales followed by Spain and France.
- Medium deal sizes are the maximum.
- Classic Cars have the maximum count followed by Vintage Cars and Motorcycles. Trains have the least count.
- The median Maximum Retail Price is approximately 100\$.
- Sales has a minimum of 500\$ to 8000\$ with the median sales being 3200\$. There are a lot of outliers in Sales.
- The median quantity ordered is 35.

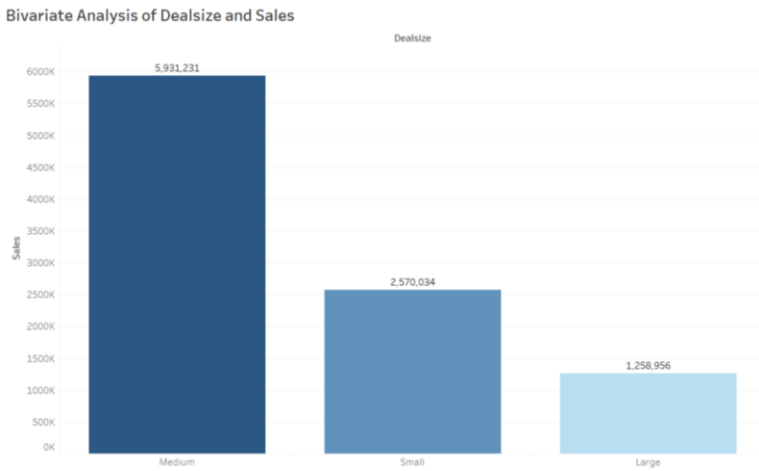
BIVARIATE ANALYSIS

Bivariate Analysis of Product and Sales

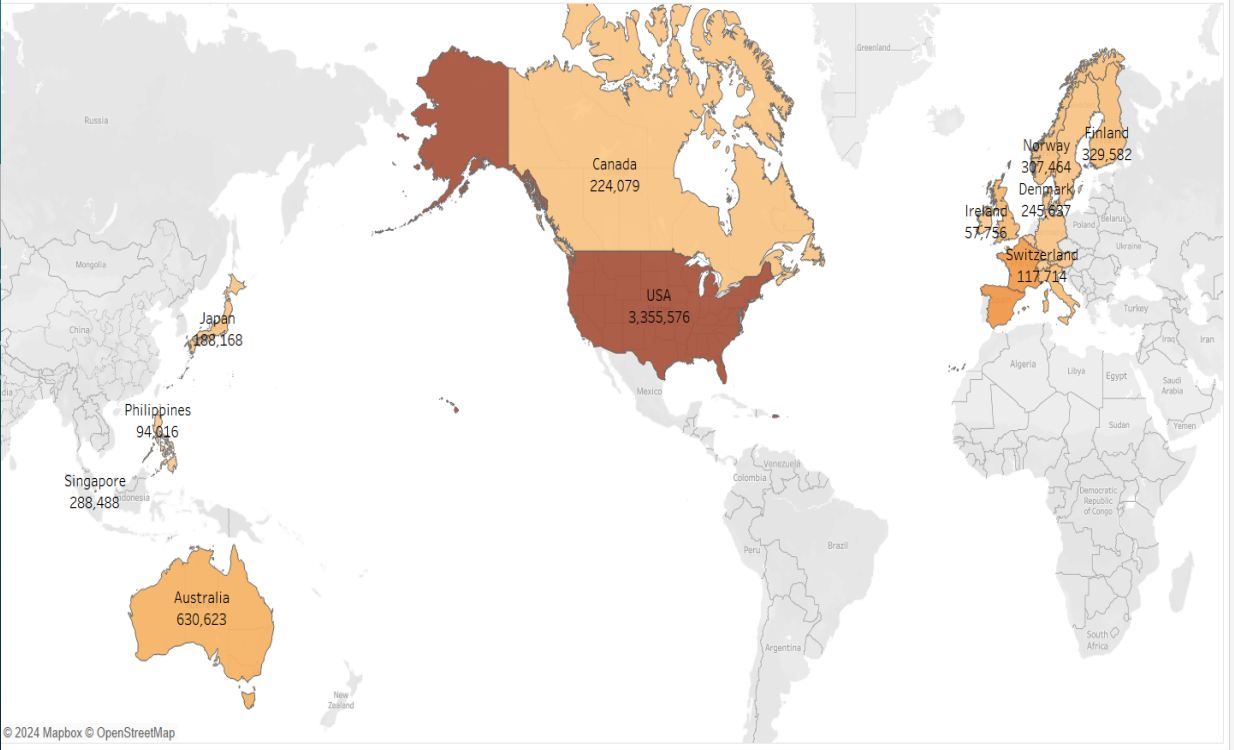


Bivariate Analysis of OrderNumber and Customer Name

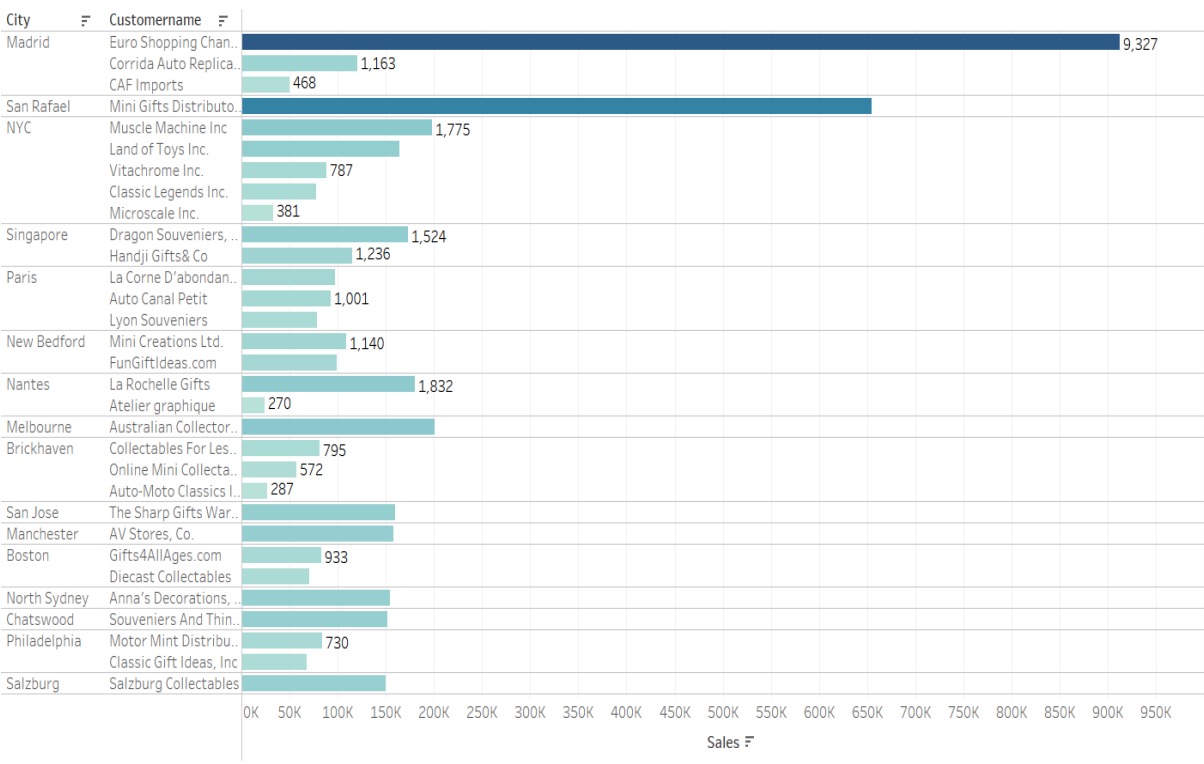




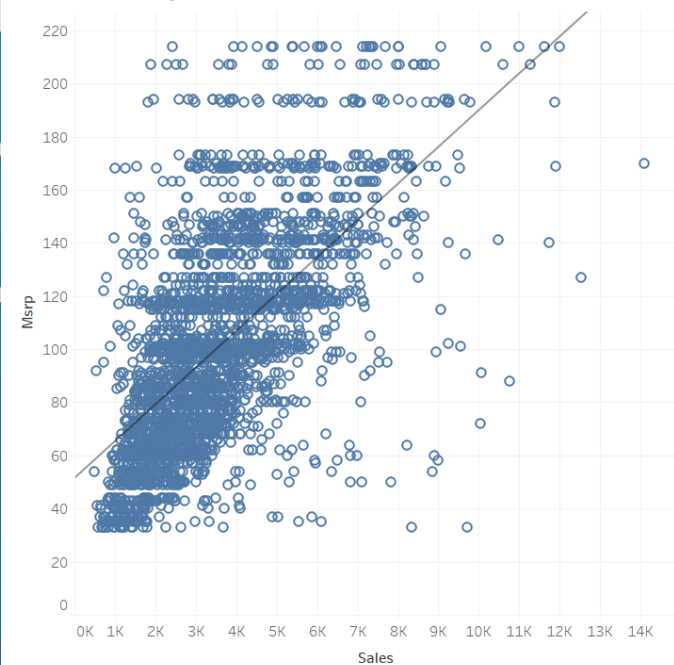
Bivariate Analysis of Country and Sales



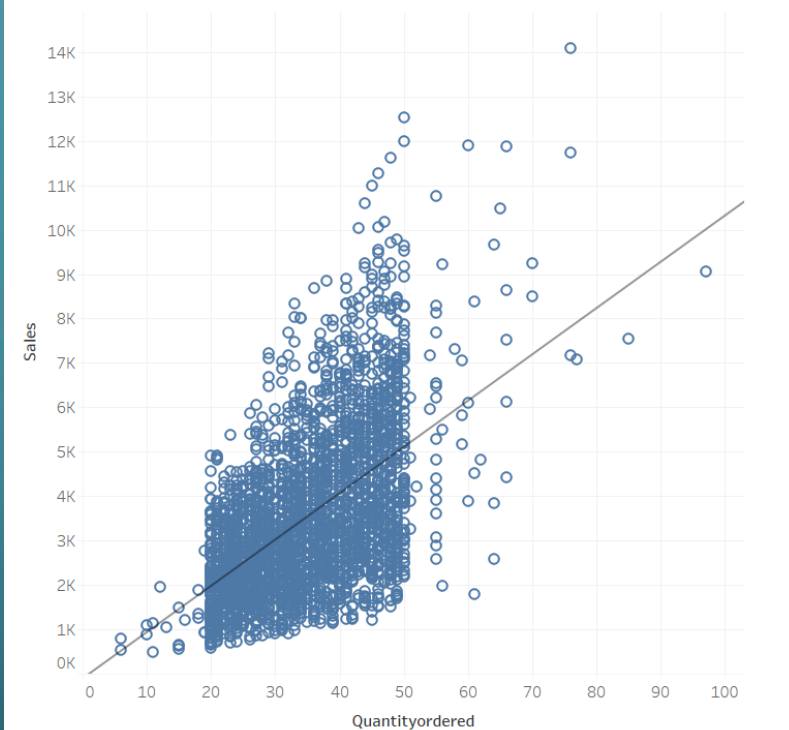
Bivariate Analysis of City and Sales



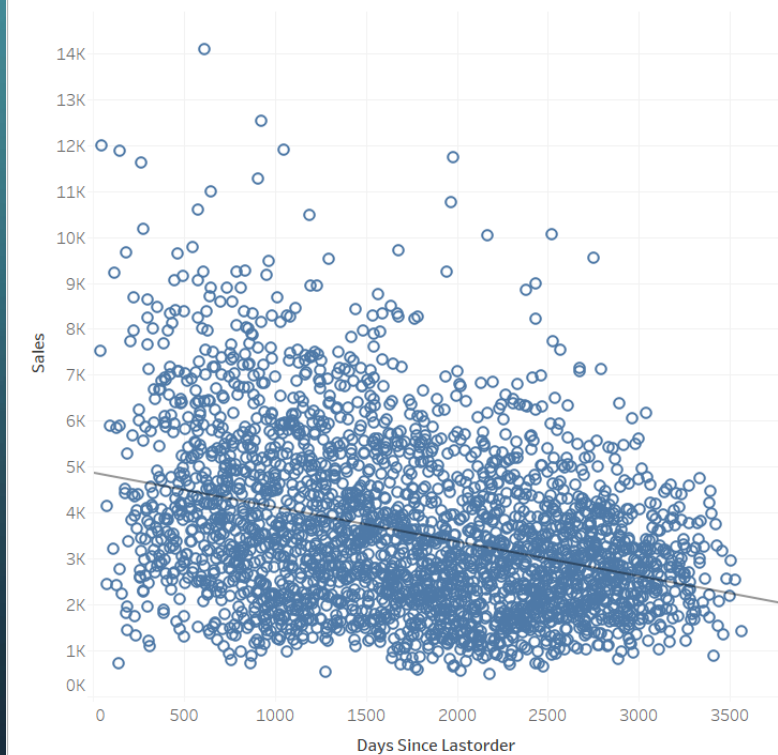
Bivariate Analysis of Sales and MSRP



Bivariate Analysis of Quantity Ordered and Sales



Bivariate Analysis of Sales and Days since Lastorder

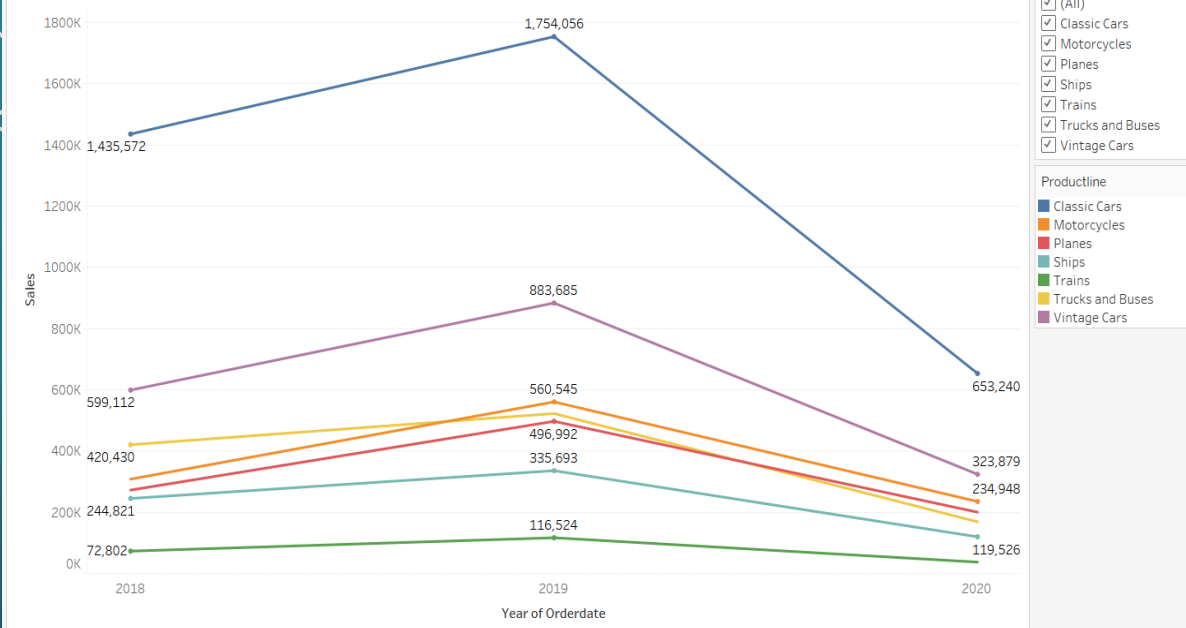


INFERENCES FROM BIVARIATE ANALYSIS

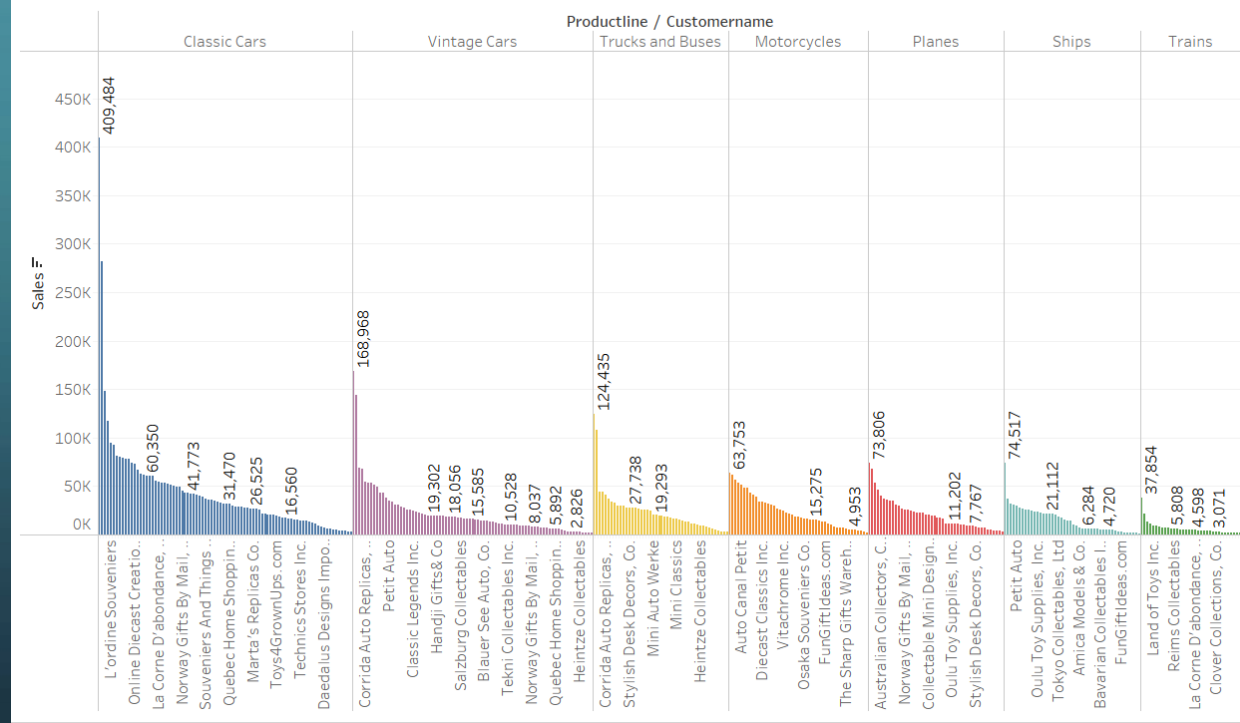
- Classic Cars have maximum sales with the top customer being L'ordine Souveneirs with a sale of approximately 4 lakh dollars.
- Among all the cities, Madrid shows maximum sales with the top customer in Madrid being Euro Shopping Channel. Madrid is followed by San Rafael and NYC.
- USA tops among all countries in sales followed by Spain.
- Sales increases exponentially with increase in MSRP, Quantity ordered as well as PriceEach.
- Sales decreases with increase in Days since last ordered.

MULTIVARIATE ANALYSIS

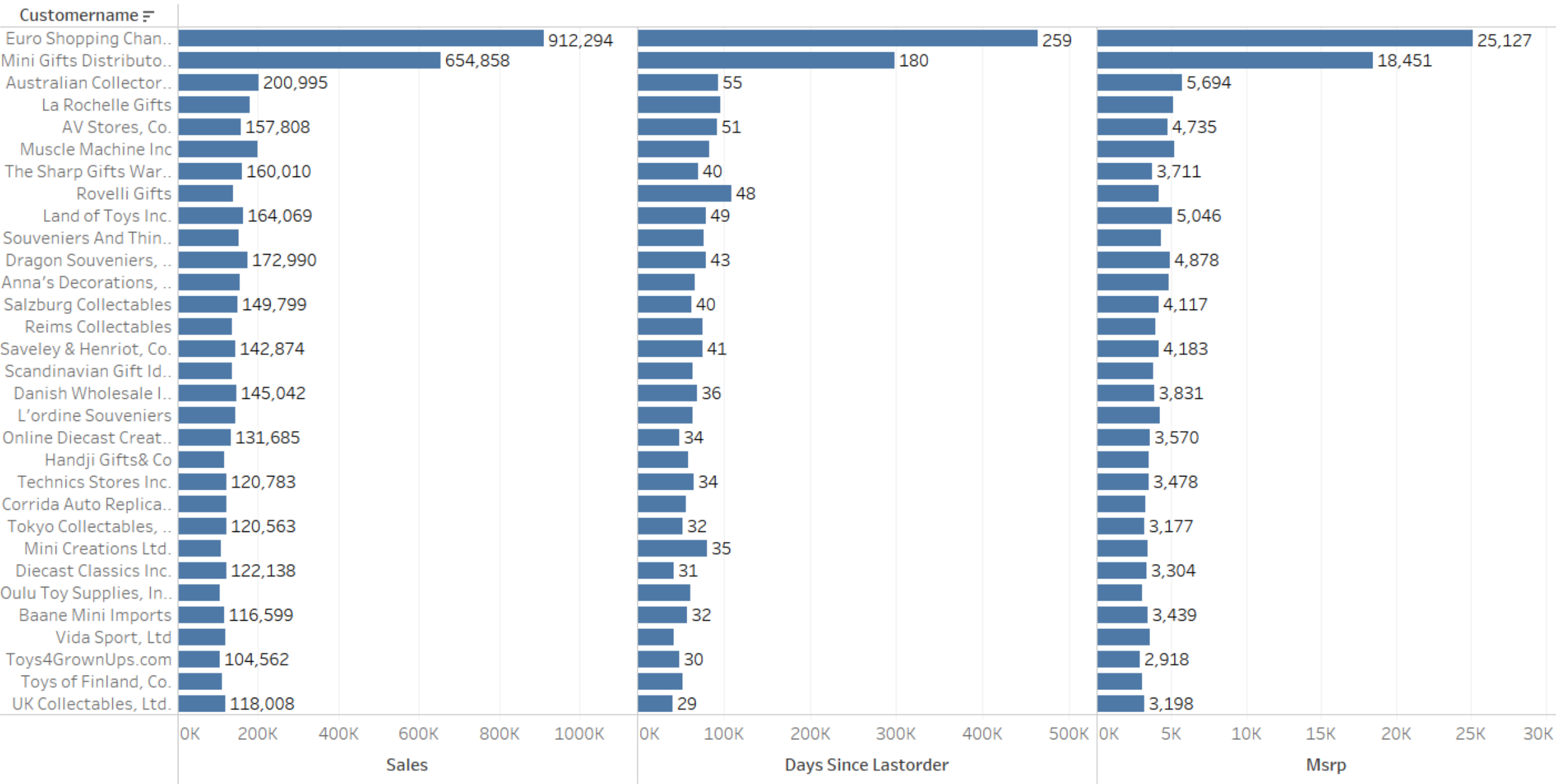
Multivariate Analysis of Yearly Sales of Products



Multivariate Analysis of Product and Sales



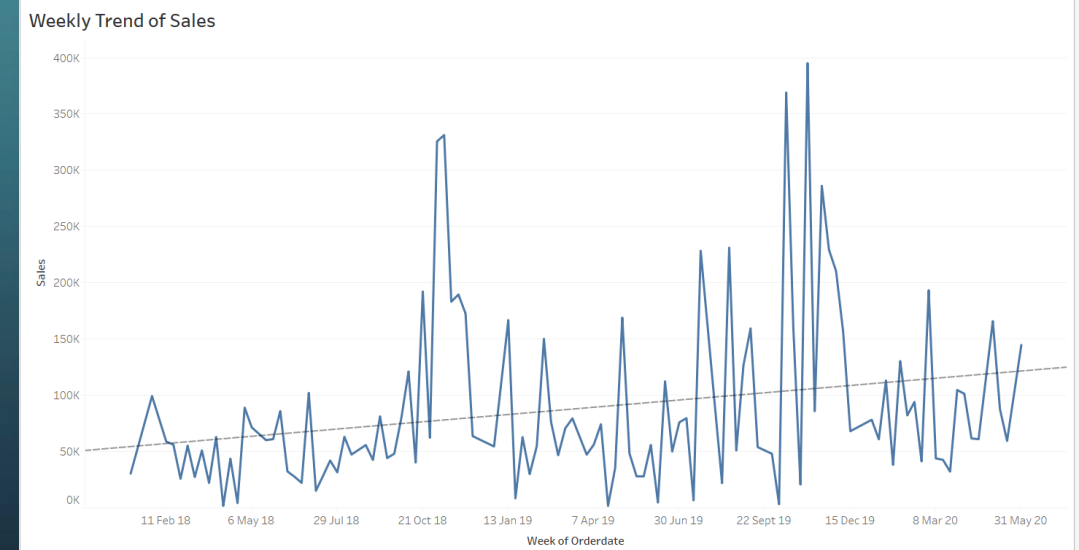
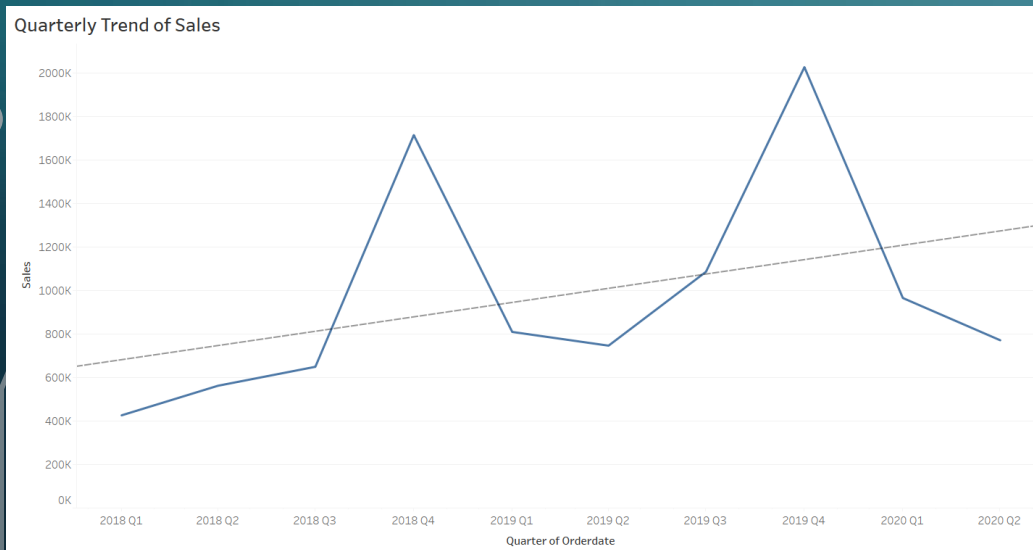
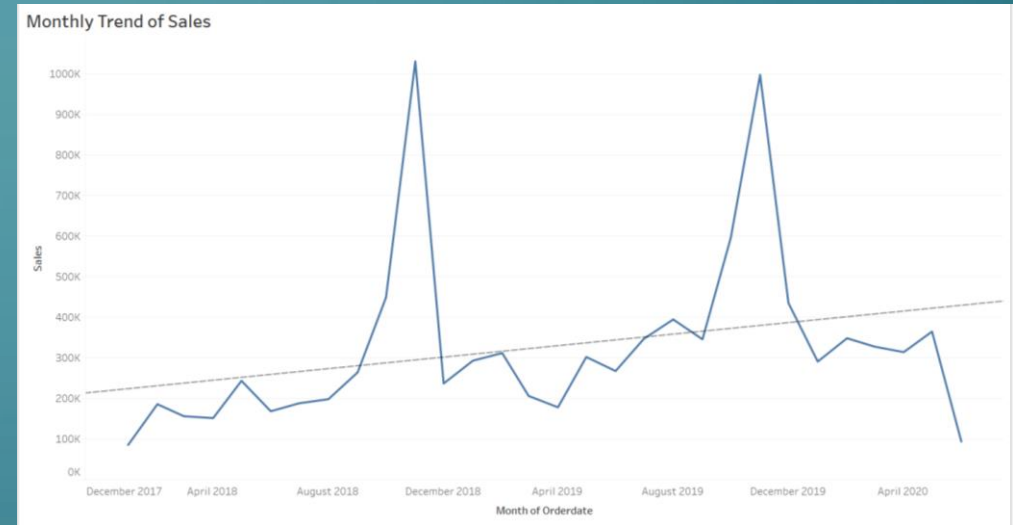
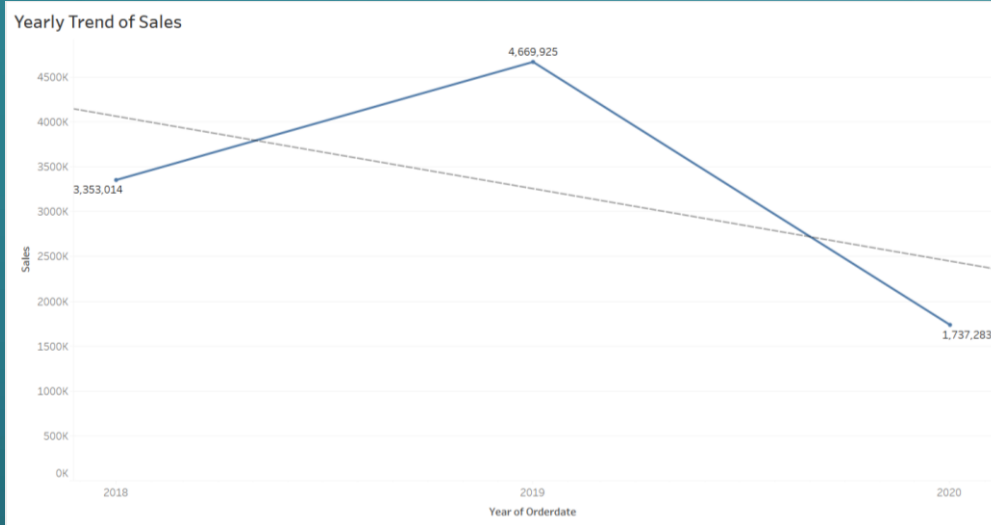
Multivariate Analysis of CustomerName with Sales, Days since Last order and MSRP



INFERENCES FROM MULTIVARIATE ANALYSIS

- L'Ordine Souveniers is the top customer in Classic Cars segment followed by Corrida Auto Replicas in Vintage Cars segment in sales.
- 2019 sees the maximum sales followed by 2018 and then 2020.
- Product sales shot up in 2019 but saw a severe fall in 2020, maybe due to covid times.
- The sales of Classic cars is double that of Vintage cars with the number of classic cars sales as high as 1754026.

WEEKLY, MONTHLY, QUARTERLY, YEARLY TRENDS IN SALES AND SALES ACROSS DIFFERENT CATEGORIES



YEARLY, QUARTERLY, MONTHLY & WEEKLY TREND OF SALES

- The year 2019 sees the maximum sales compared to 2018 and then takes a steep fall in 2020. It is a downward trend of sales from 2019 to 2020.
- The maximum sales is in the 4th quarter of 2019. The second peak is in the 4th quarter of 2018.
- November 2018 followed by November 2019 sees high peaks in sales.
- May-June, 2020 sees a massive dip in sales.
- The weekly trend sees a slight rise as the weeks increase from 2018 to 2020.
- The third and fourth weeks of November, 2019 see maximum sales.

The background is a dark teal gradient. In the corners, there are white line-art illustrations of circuit boards or neural networks, with lines connecting to small circles.

CUSTOMER SEGMENTATION USING RFM ANALYSIS (4 SEGMENTS)

WHAT IS RFM?

RFM stands for Recency, Frequency and Monetary value. It is a marketing analysis method that helps businesses segment customers and target them behaviorally.

RECENTY: How recently a customer made a purchase or interacted with the brand?

FREQUENCY: How often a customer shops or engages with a brand during a specific period?

MONETARY: How much a customer spends on the brand's products and services?

RFM uses objective, numerical scales to create a concise and high-level depiction of customers. It is simple and intuitive.

WHAT ALL PARAMETERS USED AND ASSUMPTIONS MADE?

• PARAMETERS USED:

-
- Excel Reader
- Math Formula
- Data & Time Difference
- GroupBy
- Auto-Binner
- Table Creator
- Value Lookup(Cell Replacer)
- Excel Writer

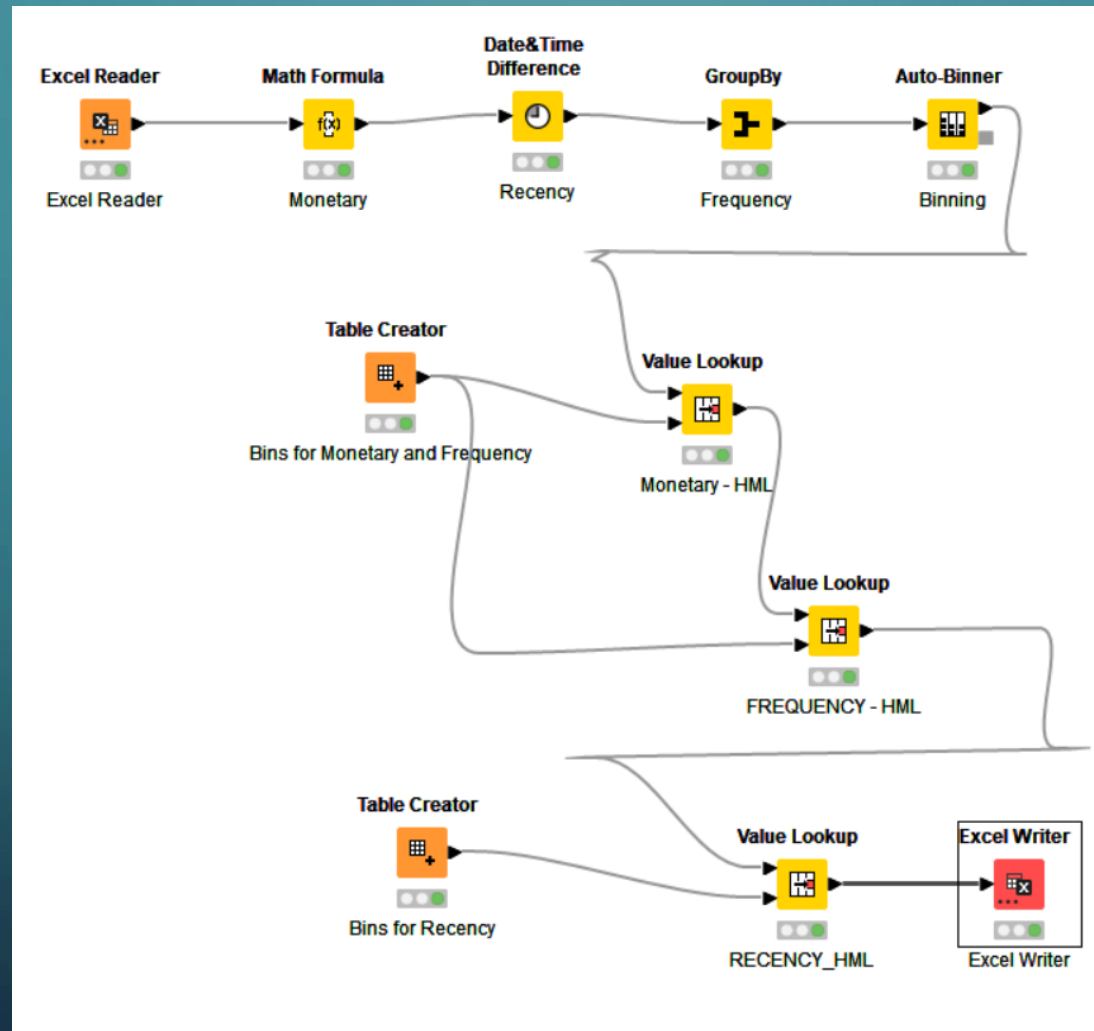
ASSUMPTIONS MADE:

The RFM model is based on the assumption that customers who have purchased recently have placed more orders and who spend more money are likely to be more valuable and are more likely to shop again.

A customer who scores a 5 for all variables – Recency, Frequency and Monetary, is your best customer while a customer who scores a 1 for all variables is your least priority customer.

SHOWCASE THE KNIME WORKFLOW IMAGE

- KNIME is used for RFM Analysis. Here is the KNIME workflow.





• CALCULATIONS USED:

- RECENCY: DATEDIFF(30-05-2020 – [ORDER DATE])
- FREQUENCY: GROUPBY done using CUSTOMERNAME and finding the number of orders per customer.
- MONETARY: [SALES]

• BINS CLASSIFIED AS:


- | | |
|--------------|-------------|
| • 0 – 0.25 | LOW |
| • 0.25 – 0.5 | LOW MEDIUM |
| • 0.5-0.75 | HIGH MEDIUM |
| • 0.75 – 1 | HIGH |
- 

WHAT RESULTS ARE THERE IN THE OUTPUT TABLE HEAD?

CUSTOMERNAME	ORDERNUMBER	QUANTITYORDERED	COUNTRY	CONTACTLASTNAME	CONTACTFIRSTNAME	DEALSIZE	MONETARY	RECENCY	FREQUENCY [Binned]	MONETARY [Binned]	RECENCY [Binned]
AV Stores, Co.	3	34.8627451	UK	Ashworth	Victoria	Medium	157807.81	195	LM	H	LM
Alpha Cognac	3	34.35	France	Roulet	Annette	Medium	70488.44	63	LM	L	H
Amica Models & Co.	2	32.42307692	Italy	Accorti	Paolo	Large	94117.26	264	L	HM	L
Anna's Decorations, Ltd	4	31.93478261	Australia	O'Hara	Anna	Small	153996.13	82	H	H	HM
Atelier graphique	3	38.57142857	France	Schmitt	Carine	Medium	24179.96	187	LM	L	LM
Australian Collectables, Ltd	3	30.65217391	Australia	Connery	Sean	Medium	64591.46	21	LM	L	H
Australian Collectors, Co.	5	35.01818182	Australia	Ferguson	Peter	Medium	200995.41	183	H	H	HM
Australian Gift Network, Co	3	36.33333333	Australia	Calaghan	Tony	Large	59469.12	118	LM	L	HM
Auto Assoc. & Cie.	2	35.38888889	France	Tonini	Daniel	Large	64834.32	232	L	L	L
Auto Canal Petit	3	37.07407407	France	Perrier	Dominique	Medium	93170.66	53	LM	HM	H
Auto-Moto Classics Inc.	3	35.875	USA	Taylor	Leslie	Medium	26479.26	179	LM	L	HM
Baane Mini Imports	4	33.8125	Norway	Bergulfsen	Jonas	Medium	116599.19	207	H	HM	LM
Bavarian Collectables Imports, Co.	1	28.64285714	Germany	Donnermeyer	Michael	Medium	34993.92	258	L	L	L
Blauer See Auto, Co.	4	36.86363636	Germany	Keitel	Roland	Medium	85171.59	207	H	LM	LM
Boards & Toys Co.	2	34	USA	Young	Leslie	Medium	9129.35	112	L	L	HM
CAF Imports	2	36	Spain	Fernandez	Jesus	Large	49642.05	438	L	L	L
Cambridge Collectables Co.	2	32.45454545	USA	Tseng	Kyung	Medium	36163.62	388	L	L	L
Canadian Gift Exchange Network	2	31.95454545	Canada	Tannamuri	Yoshi	Large	75238.92	221	L	LM	LM
Classic Gift Ideas, Inc	2	31.80952381	USA	Cervantes	Francisca	Medium	67506.97	229	L	L	LM
Classic Legends Inc.	3	36	USA	Hernandez	Maria	Medium	77795.2	191	LM	LM	LM
Clover Collections, Co.	2	30.625	Ireland	Cassidy	Dean	Large	57756.43	257	L	L	L
Collectable Mini Designs Co.	2	38.16	USA	Thompson	Valarie	Medium	87489.23	459	L	LM	L
Collectables For Less Inc.	3	33.125	USA	Nelson	Allen	Medium	81577.98	131	LM	LM	HM
Corrida Auto Replicas, Ltd	3	36.34375	Spain	Sommer	Martin	Large	120615.28	211	LM	H	LM

INFERENCES FROM RFM ANALYSIS & IDENTIFIED SEGMENTS

TYPE OF CUSTOMERS	RECENCY	FREQUENCY	MONETARY				TOTAL	
			H	HM	LM	L		
ACTIVE	H	H	9	2			11	23
		HM						
		LM	1	3	6	2	12	
		L						
AT RISK	HM	H	5			1	6	22
		HM						
		LM	1	7	3	4	15	
		L				1	1	
	LM	H	1	1	2		4	22
		HM						
		LM	4	3	3	1	11	
		L		2	3	2	7	
INACTIVE	L	H						22
		HM						
		LM	1	1			2	
		L		3	5	12	20	

- 
- A decorative graphic consisting of white lines and circles, resembling a circuit board or a stylized tree, is positioned on the left side of the slide. The lines branch out from the top and bottom, with small circles at the ends of some branches.
- The top green belt is the most active customers.
 - 9 customers are their best customers and they have the highest recency, frequency and monetary.
 - 23 customers are active customers. Out of the 23, 12 are the silver customers and a little bit of follow-up can make them into gold customers.
 - 44 customers are at risk and can be on the verge of churning. There must be strategies in place to bring these customers back.
 - 22 customers are totally inactive with the lowest recency. The 12 customers with the lowest recency, frequency and monetary value do not add anything to the sales.

WHO ARE YOUR BEST CUSTOMERS?

- 1. EURO SHOPPING CHANNEL
- 2. MINI GIFTS DISTRIBUTORS LTD.
- 3. SOUVENIERS AND THINGS Co.
- 4. LA ROCHELLE GIFTS
- 5. THE SHARP GIFTS WAREHOUSE

CUSTOMERNAME	ORDERNUMB	QUANTITYORDERI	COUNTR	CONTACTLASTNAM	CONTACTFIRSTNAM	DEALSI	MONETA	RECEN	FREQUENCY [Binne	MONETARY [Binne	REGENCY [Binne
Danish Wholesale Imports	5	36.52777778	Denmark	Petersen	Jytte	Large	145041.6	45	H	H	H
Diecast Classics Inc.	4	35.83870968	USA	Yu	Kyung	Medium	122138.14	0	H	H	H
Euro Shopping Channel	26	36.01158301	Spain	Freyre	Diego	Large	912294.11	-1	H	H	H
La Rochelle Gifts	4	34.56603774	France	Labruno	Janine	Medium	180124.9	-1	H	H	H
Mini Gifts Distributors Ltd.	17	35.36666667	USA	Nelson	Valarie	Large	654858.06	1	H	H	H
Reims Collectables	5	34.95121951	France	Henriot	Paul	Small	135042.94	61	H	H	H
Salzburg Collectables	4	36.05	Austria	Pipps	Georg	Large	149798.63	13	H	H	H
Souvenirs And Things Co.	4	34.80434783	Australia	Huxley	Adrian	Small	151570.98	1	H	H	H
The Sharp Gifts Warehouse	4	41.4	USA	Frick	Sue	Large	160010.27	38	H	H	H

WHICH CUSTOMERS ARE ON THE VERGE OF CHURNING?

- 1. ALPHA COGNAC
- 2. AUSTRALIAN COLLECTABLES, Ltd
- 3. Quebec Home Shopping Network
- 4. Lyon Souveniers
- 5. Mini Caravy

CUSTOMERNAME	ORDERNUMB	QUANTITYORDERI	COUNTF	CONTACTLASTNAM	CONTACTFIRSTNAM	DEALSI	MONETA	RECENCY	FREQUENCY [Binne	MONETARY [Binne	RECENCY [Binne
Alpha Cognac	3	34.35	France	Roulet	Annette	Medium	70488.44	63	LM	L	H
Australian Collectables, Ltd	3	30.65217391	Australia	Connery	Sean	Medium	64591.46	21	LM	L	H

CUSTOMERNAME	ORDERNUMB	QUANTITYORDERI	COUNTF	CONTACTLASTNAM	CONTACTFIRSTNAM	DEALSI	MONETA	RECENCY	FREQUENCY [Binne	MONETARY [Binne	RECENCY [Binne
Quebec Home Shopping Network	3	32.59090909	Canada	Fresnisre	Jean	Medium	74204.79	29	LM	LM	H
Petit Auto	3	31.84	Belgium	Dewey	Catherine	Medium	74972.52	0	LM	LM	H
Lyon Souveniers	3	34.2	France	Da Cunha	Daniel	Medium	78570.34	74	LM	LM	H
Mini Caravy	3	41	France	Citeaux	Frederique	Large	80438.48	46	LM	LM	H
Gifts4AllAges.com	3	35.88461538	USA	Yoshido	Juri	Medium	83209.88	24	LM	LM	H
Tekni Collectables Inc.	3	43.14285714	USA	Brown	William	Medium	83228.19	57	LM	LM	H

WHO ARE YOUR LOST CUSTOMERS ?

- Bavarian Collectables Imports, Co.
- 2. Double Decker Gift Stores, Ltd
- 3. Cambridge Collectables Co.
- 4. West Coast Collectables Co.
- 5. CAF Imports

CUSTOMERNAME	ORDERNUMB	QUANTITYORDERI	COUNTR	CONTACTLASTNAM	CONTACTFIRSTNAM	DEALS	MONETA	RECENCY	FREQUENCY [Binne	MONETARY [Binne	RECENCY [Binne
Bavarian Collectables Imports, Co.	1	28.64285714	Germany	Donnermeyer	Michael	Medium	34993.92	258	L	L	L
Double Decker Gift Stores, Ltd	2	29.75	UK	Hardy	Thomas	Medium	36019.04	494	L	L	L
Cambridge Collectables Co.	2	32.45454545	USA	Tseng	Kyung	Medium	36163.62	388	L	L	L
West Coast Collectables Co.	2	39.30769231	USA	Thompson	Steve	Medium	46084.64	487	L	L	L
CAF Imports	2	36	Spain	Fernandez	Jesus	Large	49642.05	438	L	L	L
Signal Collectibles Ltd.	2	34.26666667	USA	Taylor	Sue	Medium	50218.51	475	L	L	L
Iberia Gift Imports, Corp.	2	39.26666667	Spain	Roel	Jose Pedro	Medium	54723.62	237	L	L	L
Online Mini Collectables	2	38.13333333	USA	Barajas	Miguel	Large	57197.96	263	L	L	L
Clover Collections, Co.	2	30.625	Ireland	Cassidy	Dean	Large	57756.43	257	L	L	L
Auto Assoc. & Cie.	2	35.38888889	France	Tonini	Daniel	Large	64834.32	232	L	L	L
Osaka Souvenirs Co.	2	34.6	Japan	Kentary	Mory	Medium	67605.07	413	L	L	L
Daedalus Designs Imports	2	34.95	France	Rance	Martine	Small	69052.41	464	L	L	L

WHO ARE YOUR LOYAL CUSTOMERS ?

- 1. HANDJI GIFTS AND CO.
- 2. TOKYO COLLECTABLES LTD.
- 3. AUTO CANAL PETIT
- 4. GIFT DEPOT INC.
- 5. UK COLLECTABLES LTD.

CUSTOMERNAME	ORDERNUMB	QUANTITYORDERI	COUNTF	CONTACTLASTNAM	CONTACTFIRSTNAM	DEALSI	MONETA	REGENCY	FREQUENCY [Binne	MONETARY [Binne	REGENCY [Binne
Auto Canal Petit	3	37.07407407	France	Perrier	Dominique	Medium	93170.66	53	LM	HM	H
Gift Depot Inc.	3	36.12	USA	King	Julie	Medium	101894.79	25	LM	HM	H
Handji Gifts& Co	4	34.33333333	Singapore	Victorino	Wendy	Large	115498.73	37	H	HM	H
Tokyo Collectables, Ltd	4	35.9375	Japan	Shimamura	Akiko	Large	120562.74	38	H	HM	H
UK Collectables, Ltd.	3	36.06896552	UK	Devon	Elizabeth	Small	118008.27	52	LM	HM	H



THE END

