

CUSTOMER BUYING PATTERNS – RFM ANALYSIS

PGP-DSBA

KARTHICK RAJ S



2 TABLE OF CONTENTS

S.No	Title	Slide No.
1	Part A - Problem statement	<u>3</u>
2	Agenda & Executive Summary of the data	<u>4</u>
3	About Data	<u>5</u>
4	Descriptive Statistics	<u>6</u>
5	Exploratory Analysis and Inferences	<u>8</u>
6	KNIME workflow	<u>16</u>
7	RFM analysis	<u>17</u>
8	Conclusion	<u>22</u>
9	Appendix	<u>23</u>

3 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.

4 AGENDA & EXECUTIVE SUMMARY OF THE DATA

- Agenda:
 - ✓ Understand the buying behaviour of the customers.
 - ✓ Provide insights about the customer and their segments.
- Executive Summary:
 - ✓ The data is from an automobile company having the sales, geographic and other details of their products & buyers over the past three years.

5 ABOUT DATA

- The Data has 2727 Rows and 20 Columns.
- It has 6 Numerical, 13 Categorical and 1 Date fields.
- The Data Explorer Node in the KNIME has excluded the Date field for Descriptive Statistics.

6 DESCRIPTIVE STATISTICS - NUMERICAL DATA

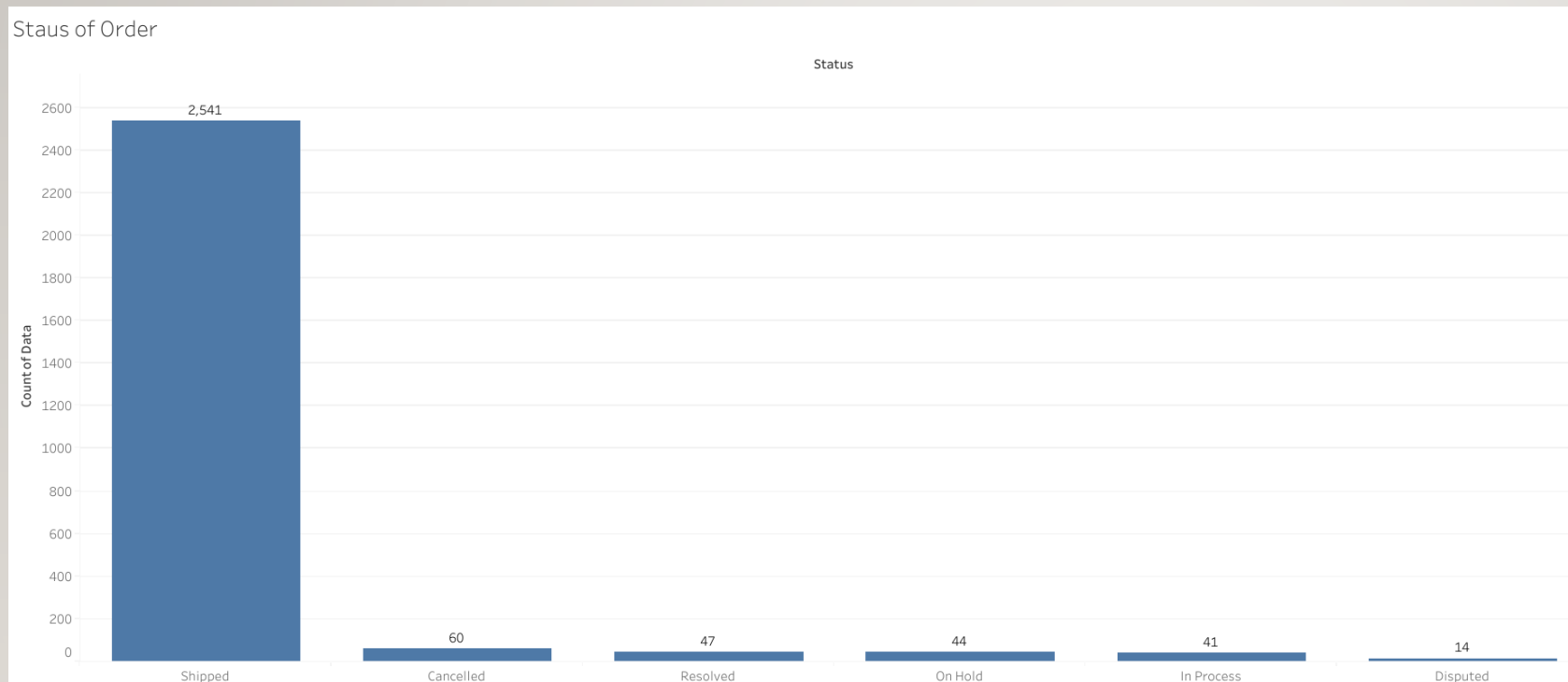
Column	Exclude Column	Minimum	Maximum	Mean	Standard Deviation	Variance	Skewness
+ QUANTITYORDERED	<input type="checkbox"/>	6	97	35.103	9.762	95.299	0.369
+ PRICEEACH	<input type="checkbox"/>	26.880	252.870	101.099	42.043	1767.576	0.697
+ ORDERLINENUMBER	<input type="checkbox"/>	1	18	6.491	4.231	17.897	0.575
+ SALES	<input type="checkbox"/>	482.130	14082.800	3553.048	1838.954	3381751.448	1.156
+ DAYS_SINCE_LASTORDER	<input type="checkbox"/>	42	3562	1757.086	819.281	671220.663	-0.003
+ MSRP	<input type="checkbox"/>	33	214	100.692	40.115	1609.197	0.576

- The Above table shows the descriptive statistics for Numerical data.
- The Maximum sales recorded in the transaction for three years is \$14,082.80.

7 DESCRIPTIVE STATISTICS - CATEGORICAL DATA

Name	Type	Missing Values	Unique Values
ORDERNUMBER	String	0	298
STATUS	String	0	6
PRODUCTLINE	String	0	7
PRODUCTCODE	String	0	109
CUSTOMERNAME	String	0	89
PHONE	String	0	88
ADDRESSLINE1	String	0	89
CITY	String	0	71
POSTALCODE	String	0	73
COUNTRY	String	0	19
CONTACTLASTNAME	String	0	76
CONTACTFIRSTNAME	String	0	72
DEALSIZE	String	0	3

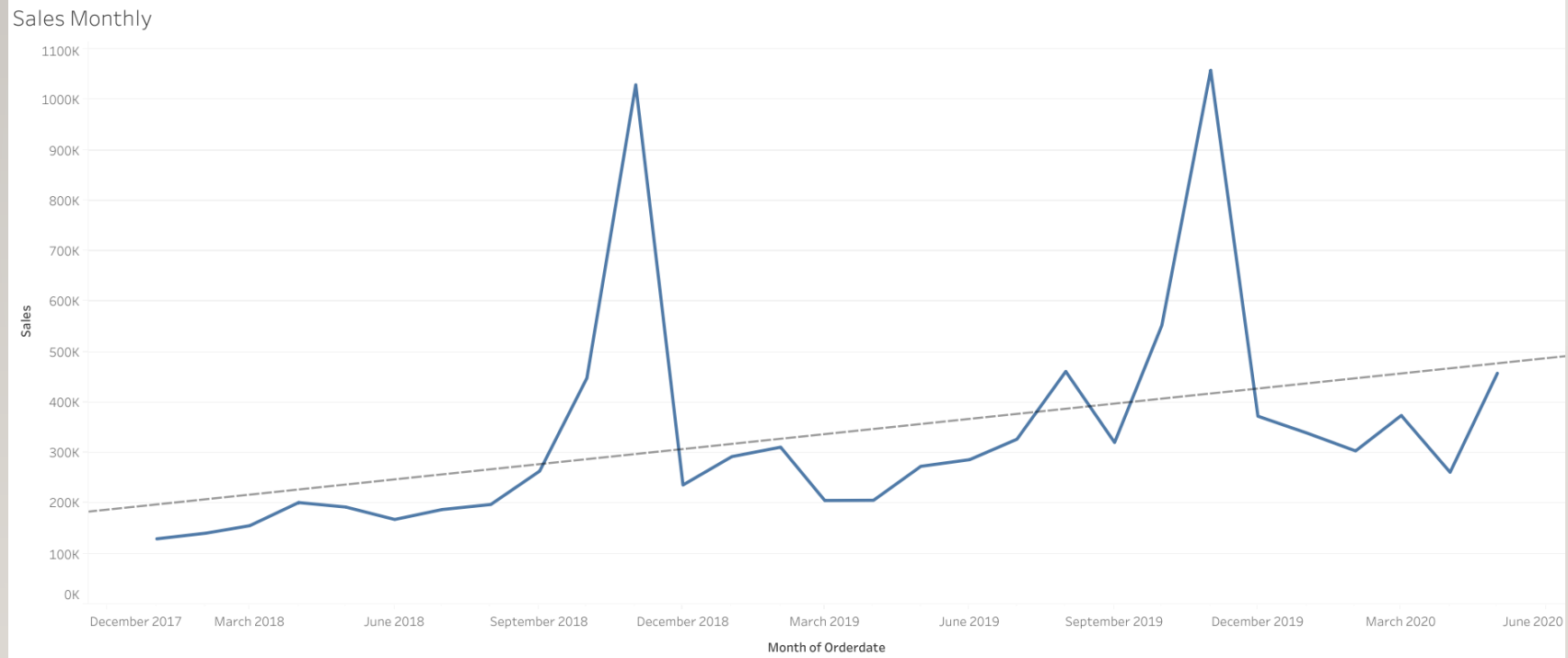
8 EXPLORATORY ANALYSIS AND INFERENCES



- Most of the orders are shipped which is a good sign for the company.

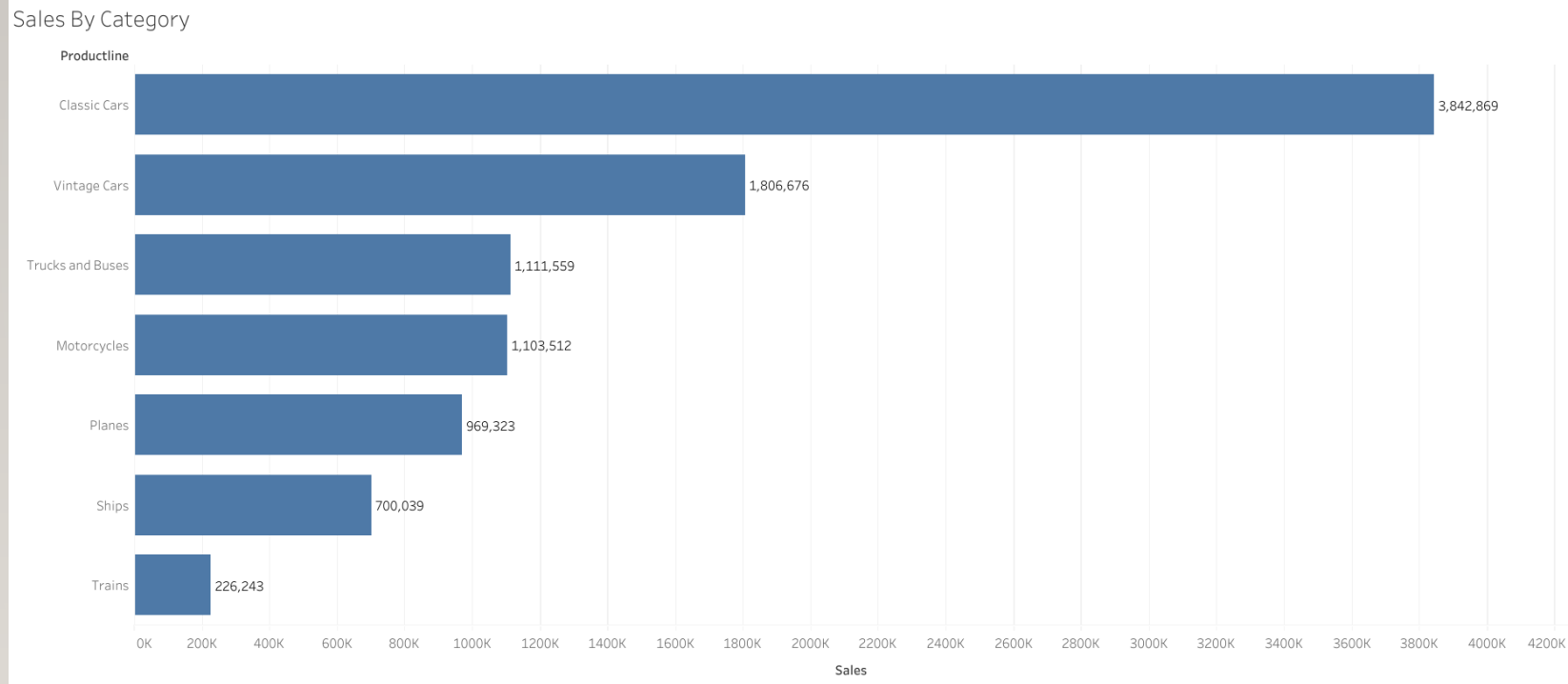
Status	Count	%
Shipped	2541	92.50%
Cancelled	60	2.18%
Resolved	47	1.71%
On Hold	44	1.60%
In Process	41	1.49%
Disputed	14	0.51%
Total	2747	100.00%

9 MONTHLY SALES



- There is an upward trend in the sales.
- The seasonality is present as the sales increases in the month of November.

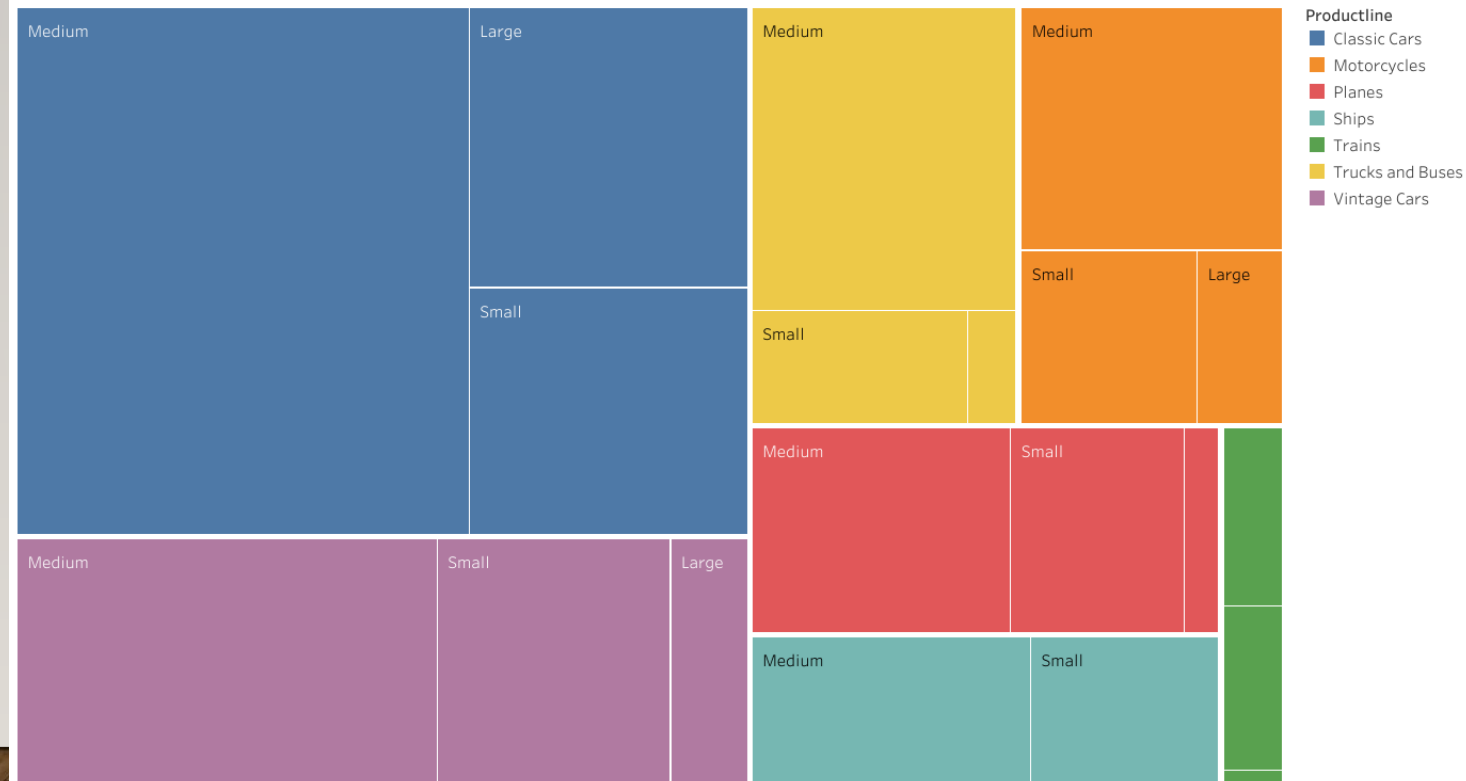
10 SALES BY CATEGORY



- The Most sales recorded is the Classic cars Category with \$ 3.8 M.
- The least sales recorded is by Trains.

SALES BY DEAL SIZE & CATEGORY

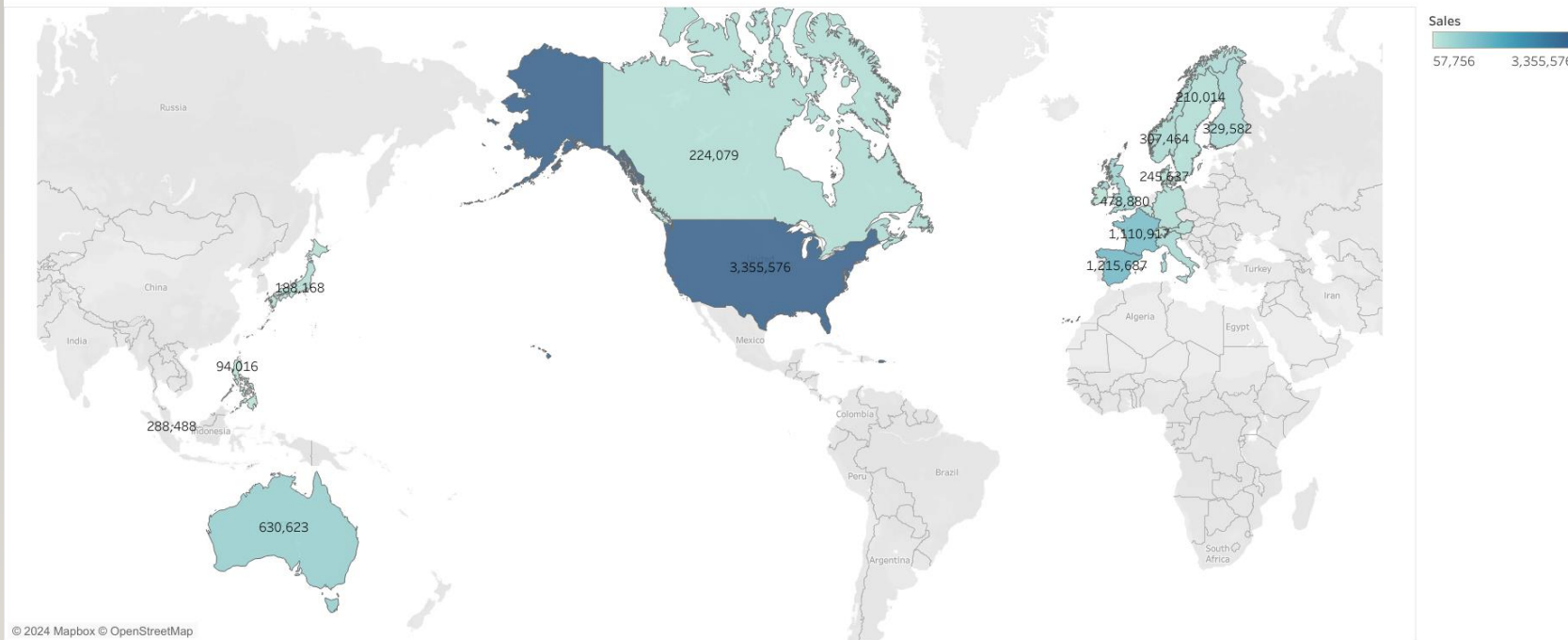
Sales by Deal size and Category



- All the category has most of the sales in Medium Deal size then small or large deal size.

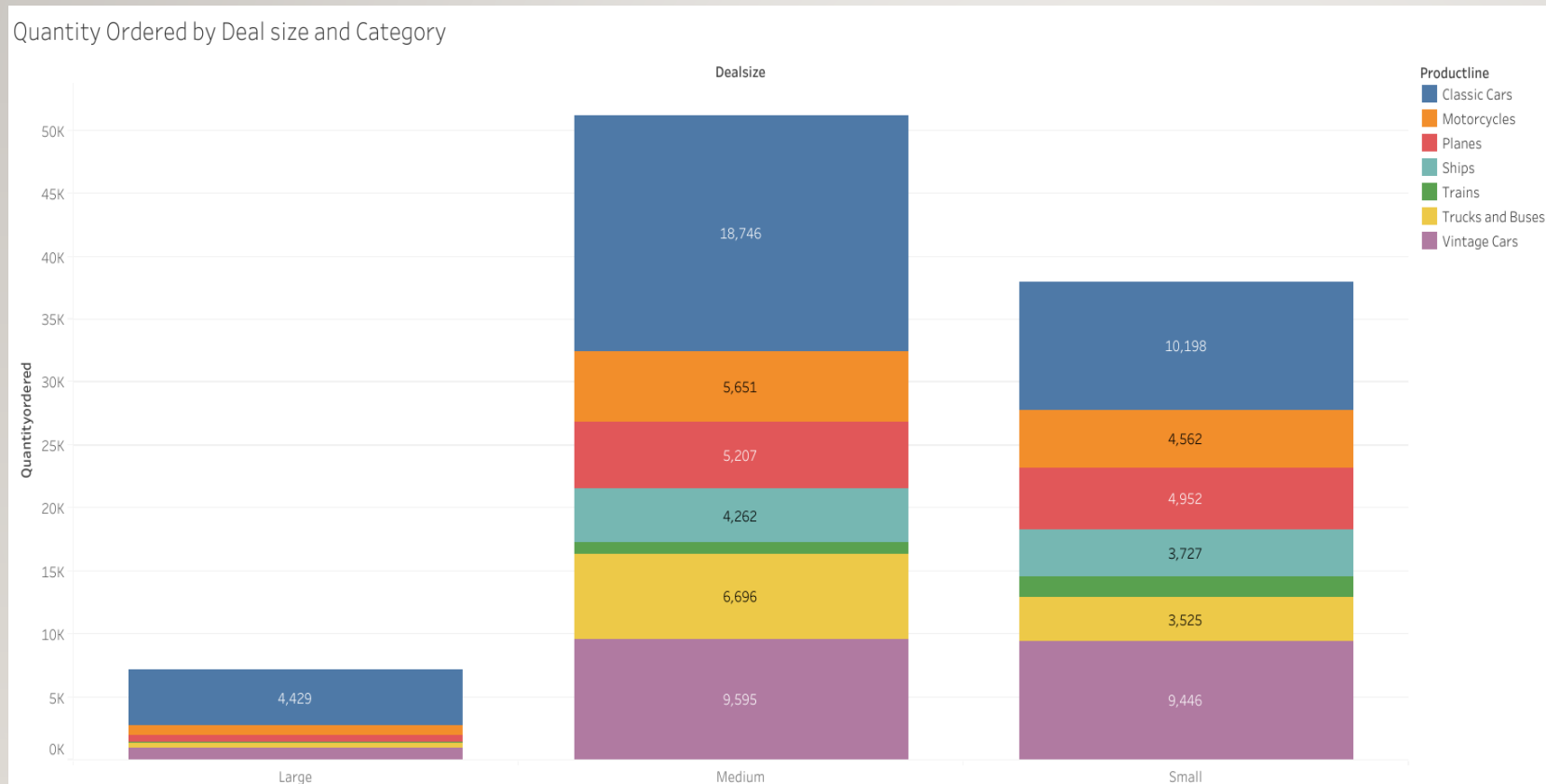
12 COUNTRY WISE SALES

Country Wise Sales



- More than half of the sales for the company is coming from USA, Spain and France Customers.
- Geographical wise this segment is the highest sales for the company.
- In USA, San Rafael and NYC are the highest sales making cities.

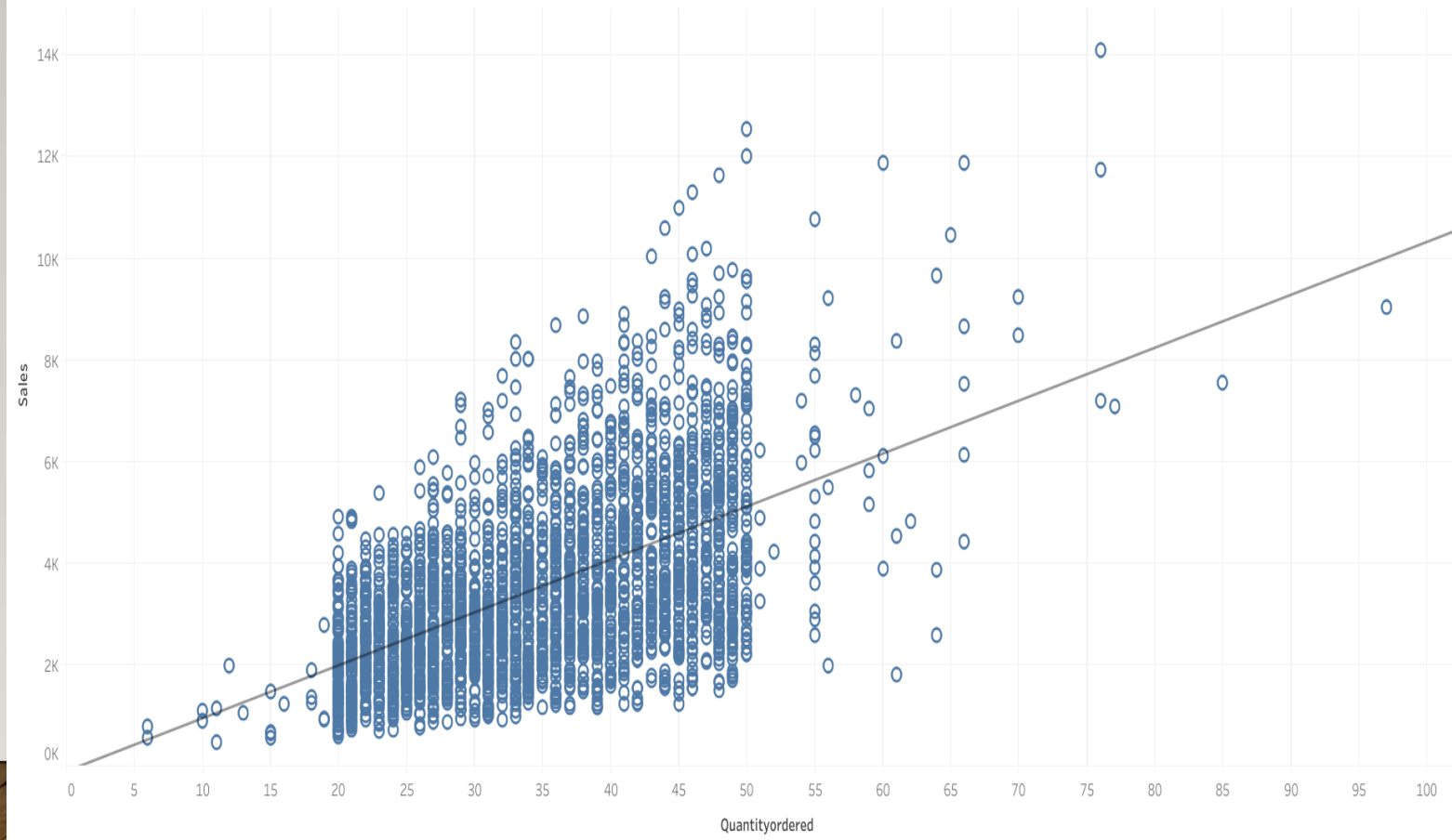
I3 QUANTITY ORDERED BY DEAL SIZE & CATEGORY



- The Highest quantity ordered is by Medium deal size in size category.
- In terms of Product line category, Classic cars has the highest ordered size.

I4 SCATTER PLOT – SALES VS QUANTITY

Scatter Plot - Sales Vs Quantity



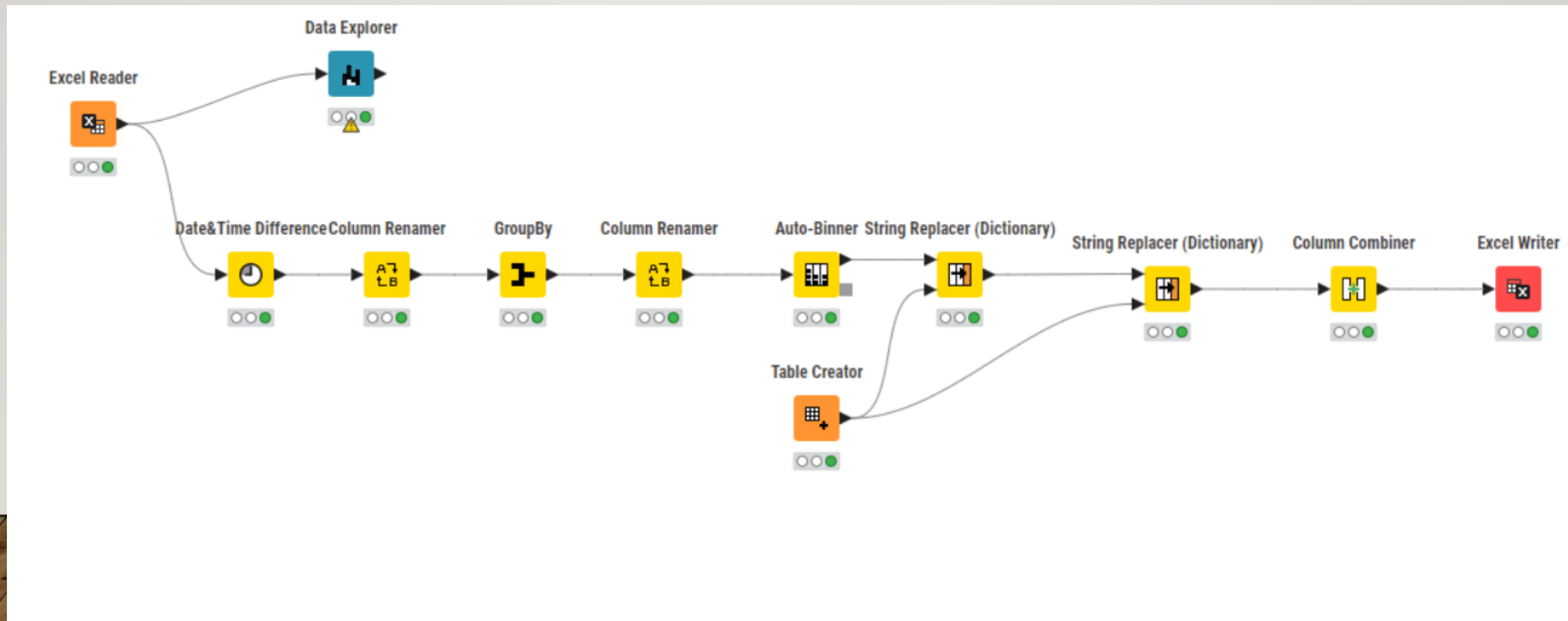
- There is an positive trend relationship between the sales and quantity.
- As the Quantity increases, sales also increases.
- Most of the Quantity ordered is between the range of 20 to 50.

15 SUMMARISE THE INFERENCES

- The Maximum sales recorded in the transaction for three years is \$14,082.80.
- There is an upward trend in the sales and seasonality is present as the sales increases in the month of November.
- The Highest and Lowest sales recorded is for Classic cars Category and Trains respectively.
- All the category has most of the sales in Medium Deal size then small or large deal size.
- Geographical wise USA, Spain and France Customer's are the highest sales making segments for the company.
- The Highest quantity ordered is by Medium deal size in size category and interms of Product line category, Classic cars has the highest ordered size.
- There is an positive trend relationship between the sales and quantity, as the Quantity increases, sales also increases.

16 KNIME WORKFLOW

- This is the KNIME Workflow for the RFM Analysis.
- The Excel Output table has Customer Name, Recency & Frequency count, Monetary value and RFM Bins.



17 RFM ANALYSIS

- RFM Analysis uses the Recency, Frequency and Monetary of the customers purchase behaviour to understand the segments and their behaviours.
- For the Automobile customers, the assumptions for RFM are:
 - ✓ Splitting of Bins into three types as High (H), Medium(M) and Low(L) for Recency, Frequency and Monetary.
 - ✓ The Days since last order has been calculated with the current date instead of the one given in the dataset as it has many mistake in calculation.
 - ✓ The Dataset is grouped by Customername.
 - ✓ The RFM is used and the dataset is analysed for those 89 customers.

18 RFM SEGMENTS

		Monetary		
Recency	Frequency	H	M	L
H	H	9	2	0
	M	1	9	2
	L	0	0	0
M	H	6	3	1
	M	5	16	5
	L	0	5	3
L	H	0	0	0
	M	1	1	0
	L	0	8	12

Colour	Segments	Count
	Tier 1 Customers	15
	Tier 2 Customers	38
	Tier 3 Customers	21
	Tier 4 Customers	15

Behaviour	Count	%
Loyal	21	23.60%
High Profit	22	24.72%
Most Frequent	21	23.60%
Lost	15	16.85%

19 INFERENCES FROM RFM ANALYSIS

- The Customer segment is split into 4 categories from Tier 1 to Tier 4.
- The Tier 1 customers are premium customers with loyalty and high monetary value.
- Tier 1 and Tier 2 customers gives the most amount of sales.
- Tier 2 customers have high monetary value but they are not loyal as that of premium customers.
- Tier 4 customers are the lost ones.
- Some of the tier 3 customers are going to become lost customers.

20 CUSTOMER BEHAVIOUR

- **Who are your best customers?**
 - ✓ Around 25% of the customers are giving high sales. They are Royal customers.
- **Which customers are on the verge of churning?**
 - ✓ 23% of the customers might be on the verge of churning.
- **Who are your lost customers?**
 - ✓ 16% of the customers are lost.
- **Who are your loyal customers?**
 - ✓ 23% of the customers are loyal customers.

21 CUSTOMER NAME BY BEHAVIOUR

Loyal	Best	Verge	Lost
Anna's Decorations, Ltd	AV Stores, Co.	Alpha Cognac	Auto Assoc. & Cie.
Australian Collectors, Co.	Anna's Decorations, Ltd	Amica Models & Co.	Bavarian Collectables Imports, Co.
Baane Mini Imports	Australian Collectors, Co.	Atelier graphique	Boards & Toys Co.
Blauer See Auto, Co.	Corrida Auto Replicas, Ltd	Australian Collectables, Ltd	CAF Imports
Danish Wholesale Imports	Danish Wholesale Imports	Australian Gift Network, Co	Cambridge Collectables Co.
Diecast Classics Inc.	Diecast Classics Inc.	Auto-Moto Classics Inc.	Classic Gift Ideas, Inc
Dragon Souvenirs, Ltd.	Dragon Souvenirs, Ltd.	Canadian Gift Exchange Network	Clover Collections, Co.
Euro Shopping Channel	Euro Shopping Channel	Collectable Mini Designs Co.	Daedalus Designs Imports
Handji Gifts& Co	L'ordine Souvenirs	Diecast Collectables	Double Decker Gift Stores, Ltd
La Rochelle Gifts	La Rochelle Gifts	Gift Ideas Corp.	Iberia Gift Imports, Corp.
Land of Toys Inc.	Land of Toys Inc.	Heintze Collectables	Microscale Inc.
Mini Gifts Distributors Ltd.	Mini Gifts Distributors Ltd.	Marta's Replicas Co.	Online Mini Collectables
Muscle Machine Inc	Muscle Machine Inc	Mini Auto Werke	Osaka Souvenirs Co.
Reims Collectables	Online Diecast Creations Co.	Mini Classics	Signal Collectibles Ltd.
Royale Belge	Reims Collectables	Norway Gifts By Mail, Co.	West Coast Collectables Co.
Salzburg Collectables	Rovelli Gifts	Royal Canadian Collectables, Ltd.	-
Souvenirs And Things Co.	Salzburg Collectables	Super Scale Inc.	-
Technics Stores Inc.	Saveley & Henriot, Co.	Toms Spezialitten, Ltd	-
The Sharp Gifts Warehouse	Scandinavian Gift Ideas	Vida Sport, Ltd	-
Tokyo Collectables, Ltd	Souvenirs And Things Co.	giftsbymail.co.uk	-
Volvo Model Replicas, Co	Technics Stores Inc.	-	-
-	The Sharp Gifts Warehouse	-	-

22 CONCLUSION

- The Royal and loyal customers can have some premium offers so they can be retained for longer periods.
- The main concern is the tier 3 customers who are on the verge of churning. They have to be retained at some cost of better strategies and offers.
- Based on the 89 customers, the lost customers % is high. It should not increase further so the tier 3 customers has to be retained.

23 APPENDIX

- Tableau Link:
 - ✓ RFM - [Tableau Link - Part A](#)